

STIC-Biotech/ChemLib

From: Li, Bao-Qun
Sent: Friday, December 23, 2005 8:25 AM
To: STIC-Biotech/ChemLib
Subject: 10761006

Please do the amino acid sequence homology and interference searches against nucleotides from 155 to 835 and from 527 to 595 of SEQ ID NO: 1. Thank you.

Bao Qun Li M.D
TC 1600
Art Unit 1648
Tel. 517-272-0904
REM, 3C18
Rm. 3D24

RECEIVED
DEC 23 2005
STIC

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 1, 2005, 23:14:00 ; Search time 464.5 Seconds
(without alignments)
6082.251 Million cell updates/sec

Title: US-10-761-006A-1

Perfect score: 5907

Sequence: 1 CTCACACAACTCCACCAAG.....CCTCAGGCCACGAGTGGA 3215

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO_spool_p/US10761006/runat_01122005_113952_26862/app_query.fasta_1.3399
-DB=A_Geneseq -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10761006@cgn.1.1.746@runat_01122005_113952_26862 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A_Geneseq_21.*
1: Geneseqp1980a.*
2: Geneseqp1990a.*
3: Geneseqp2000a.*
4: Geneseqp2001a.*
5: Geneseqp2002a.*
6: Geneseqp2003a.*
7: Geneseqp2003bs.*
8: Geneseqp2004a.*
9: Geneseqp2005a.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2927	49.6	843	3 AAY54044	Aay54044 Amino aci
2	2810	47.6	540	9 ADX40768	Adx40768 HBV polym
3	2810	47.6	540	9 ADX40767	Adx40767 HBV polym
4	2807	47.5	843	9 ADX40772	Adx40772 HBV polym
5	2789	47.2	843	9 ADX40776	Adx40776 HBV polym
6	2778	47.0	845	9 ADX40777	Adx40777 HBV polym
7	2759	46.7	845	2 AAR70065	Aar70065 Hepatitis
8	2750	46.6	843	9 ADX40775	Adx40775 HBV polym
9	2737.5	46.3	539	7 ADM68212	Adm68212 GRP94 pro

10	2734.5	46.3	842	9	ADX40771	Adx40771 HBV polym
11	2688	45.5	825	9	ADX40774	Adx40774 HBV polym
12	2678.5	45.3	827	9	ADX40773	Adx40773 HBV polym
13	2675	45.3	842	4	AAE04708	Aae04708 Hepatitis
14	2675	45.3	842	4	AAE04708	Aae04708 Hepatitis
15	2654	44.9	832	9	ADX40761	Adx40761 HBV polym
16	2651	44.9	832	9	ADX40762	Adx40762 HBV polym
17	2643	44.7	843	9	ADX40760	Adx40760 HBV polym
18	2642	44.7	832	9	ADX40763	Adx40763 HBV polym
19	2628	44.5	832	9	ADX40765	Adx40765 HBV polym
20	2608	44.2	843	3	RAY44348	Ray44348 Human hep
21	2608	44.2	845	9	ADX40758	Adx40758 HBV polym
22	2607	44.1	832	9	ADX40764	Adx40764 HBV polym
23	2607	44.1	845	9	ADX40759	Adx40759 HBV polym
24	2582	43.7	843	9	ADX40766	Adx40766 HBV polym
25	1829	31.0	373	9	ADM42965	Adm42965 ADV (adef
26	1829	31.0	373	9	ADM42965	Adm42965 ADV (adef
27	1586	26.8	344	9	AEA08851	Aea08851 Hepatitis
28	1557	26.4	344	9	AEA08850	Aea08850 Hepatitis
29	1552	26.3	303	9	ADX40769	Adx40769 HBV polym
30	1552	26.3	303	9	ADX40770	Adx40770 HBV polym
31	1496	25.3	400	3	RAY54045	Ray54045 Amino aci
32	1479	25.0	309	9	AEA08832	Aea08832 Hepatitis
33	1465	24.8	281	1	AAP60163	Aap60163 Subtype a
34	1465	24.8	281	1	AAP60560	Aap60560 Hepatitis
35	1465	24.8	281	1	AAP60617	Aap60617 C-termina
36	1462	24.8	307	9	AEA08833	Aea08833 Hepatitis
37	1461	24.7	281	1	AAP70294	Aap70294 Subtype a
38	1459	24.7	400	2	AAR93801	Aar93801 Hepatitis
39	1459	24.7	400	6	ABR55863	Abri55863 HBV S-pro
40	1459	24.7	400	8	ADN49718	Adn49718 Hepatitis
41	1459	24.7	400	8	ADU74394	Adu74394 HBV S-pro
42	1457	24.7	348	7	ADG76935	Adg76935 Hepatitis
43	1456	24.6	281	2	AAR62870	Aar62870 Hepatitis
44	1456	24.6	281	4	AMM48419	Am48419 Protein #
45	1452	24.6	389	1	AAP60794	Aap60794 Atr-type

ALIGNMENTS

RESULT 1

AAY54044

ID AAY54044 standard; protein; 843 AA.

AC AAY54044;

DT 27-MAR-2000 (first entry)

XX Amino acid sequence of a HBV DNA polymerase protein.

KW HBV, HBV surface antigen-'S'-l33 Oon strain (Met to Thr); DNA polymerase;
KW large surface antigen; core protein; transactivating X protein;
KW hepatitis vaccine; HBV infection; hepatocellular carcinoma.

OS Hepatitis B virus.

FX Key Location/Qualifiers

FT Misc-difference 194 /note= "encoded by GAG"

FT Misc-difference 195 /note= "encoded by CAA"

XX WO9966048-A1.

PD 23-DEC-1999.

PF 19-JUN-1998; 98WO-SG000046.

PR 19-JUN-1998; 98WO-SG000046.

XX (GOVE-) GOVERNMENT REPUBLIC SINGAPORE.

XX Oon CJ, Lim GK, Zhao Y, Chen WN;

XX WPI: 2000-106104/09.
 DR N-PSDB; AAZ37088.
 XX
 PT New isolated hepatitis B virus strain, useful for, e.g. treatment of
 PT hepatitis infection.
 XX
 PS Disclosure; Page 36-39; 68pp; English.
 XX
 CC The present sequence is encoded by the genome of an isolated strain of
 CC Hepatitis B virus designated human Hepatitis B virus (HBV) surface
 CC antigen-'S'-133 Oon strain (Met to Thr). The viral genome is deposited as
 CC ECCC accession numbers P97121501, P97121502 and P97121503. The nucleotide
 CC sequence was isolated from hepatocellular carcinoma (HCC). The nucleotide
 CC sequence encodes four overlapping proteins, which are a DNA polymerase, a
 CC large surface antigen, a core protein, and a transactivating X protein.
 CC The large surface antigen differs from the wild type sequence in that it
 CC contains a Thr at position 133 of the wild type sequence instead of a
 CC Met. The proteins are used to produce antibodies. The proteins,
 CC polynucleotide and antibodies can be used for detecting the novel HBV
 CC strain. The HBV polypeptides can also be used in hepatitis vaccines. The
 CC HBV novel strain polypeptides can be used to identify compounds for
 CC treating or preventing HBV infection or hepatocellular carcinoma
 XX
 SQ Sequence 843 AA;
 Alignment Scores:
 Pred. No.: 9,31e-224 Length: 843
 Score: 2927.00 Matches: 540
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 49.55% Indels: 0
 DB: 3 Gaps: 0
 US-10-761-006A-1 (1-3215) x AAY54044 (1-843)
 QY 1 CTCACACATTCACCAAGCTCTGTAGATCCAGGTCAGGGGCTATATTTCTCTGC 60
 DB 304 LeuHisAsnIleProProSerSerAlaArgSerGlnGlyGluGlyProIlePheSerCys 323
 QY 61 TGTGGCTCCAGTTCGGGAACAGTAAACCTGTTCGACTACTGCTCTCCCATATCGTC 120
 DB 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 343
 QY 121 AATCTTCTCGAGGACTGGGACCTCCACCGAATCGGAGAACACATCAGGATTCCT 180
 DB 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyHisAsnIleArgIlePro 363
 QY 181 AGGACCTCTCGTGTACAGGGGGGTTTCTCGTTGACAGAGATCCTCAATACC 240
 DB 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
 QY 241 GCAGAGTCTAGACTGTGGTGGACTTCTCAATTTCTAGGGGGAGCACCACGTCGTTC 300
 DB 384 AlaGlnSerArgLeuTrpTrpThrSerLeuAsnPhelLeuGlyAlaProThrCysSer 403
 QY 301 TGGCCAAAATTCGAGTCCCAACCTCCATCACTACCAACCTCTGTCTCCCAATTG 360
 DB 404 TrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLeu 423
 QY 361 TCTGGCTATCGCTGGATGTCTGGGGGTTTATCATATCTCTCATCTCTGCTGCT 420
 DB 424 SerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAla 443
 QY 421 ATGCCTCATCTTCTGTGTGTCTCTCGGACTACCAAGGTATGTGCGCGTTTGTCTCT 480
 DB 444 MetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSer 463
 QY 481 ACTTCCAGGAACATCAACACCGAGCGGGGCATGCAAGACTGACGACTCCTCTCA 540
 DB 464 ThrSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSer 483
 QY 541 AGGAACTCTACGTTTCCCTCTTGTGCTGTACAAACCTTCGGACGGAACTGCACCTG 600

Db 484 ArgLysLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeu 503
 QY 601 TATTCCTCCATCCATCCTGGGCTTCGCAAGATTCCTATGGAGTGGGCTCAGTCCG 660
 DB 504 TyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPro 523
 QY 661 TTTCTCTGGCTCAGTTTACTAGTGCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCAC 720
 DB 524 PheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHis 543
 QY 721 TGTTCGGCTTTCAGTTATATATGATGATGTTGGGGCGGAAGTCTGTACAAATCTT 780
 DB 544 CysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeu 563
 QY 781 GAGTCCCTTTTACCTCTATATACCAATTTTCTTTGCTTTGGGTATACATTTAAACCT 840
 DB 564 GluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPro 583
 QY 841 AATAAACCAAGCTTGGGGCTACTCCCTTAATCTCATCGGATATGTAATGGAGTTGG 900
 DB 584 AsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrp 603
 QY 901 GGTACTTTTACCGCAGGAACATATTTACTAAACTCAAGCAATGTTTTCGAAACCTGC 960
 DB 604 GlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPro 623
 QY 961 GTAAATAGACCTATTGATTGGAAAGTATGTCAAAGAAATTTGGGTCTTTTGGGCTTTGCT 1020
 DB 624 ValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAla 643
 QY 1021 GCCCTTTTACACATGTCGCTATCTGCTTGTAGTGCCTTTATATGTCATGTATACATCT 1080
 DB 644 AlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSer 663
 QY 1081 AAGCAGGCTTTCACCTTCTCGCAACTTACAGGCTTCTGTGTAAACAAATATCTGAAC 1140
 DB 664 LysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuLysGlnTyrLeuAsn 683
 QY 1141 CTTTACCCGCTTGGCCGCAACGCTCGCTCTGCAAGTGTTCGTCAGCAACCCCTC 1200
 DB 684 LeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPro 703
 QY 1201 ACTGATGGGCTTGGCCATAGCCCATAGCGCATGGCTGGAAACCTTCTGGCTCTCTCTG 1260
 DB 704 ThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetAlaGlyThrPheLeuAlaProLeu 723
 QY 1261 CGGATCCATACGCGAAGCTCTAGCAGCTTGTTCGTCGAGCGGCTGGAGCAAAA 1320
 DB 724 ProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLys 743
 QY 1321 CTTATCGGAACGCAACACTCTGTCTCTCGGAAATACACCTCTTTCATCGGCTG 1380
 DB 744 LeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeu 763
 QY 1381 CTAGGCTGTGTCGCAACTGGATCTCGCGGGAAGTCTTCTTGTCTACGTCGCTCGGCG 1440
 DB 764 LeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAla 783
 QY 1441 CTGAATCCCGGAGACCCGCTCGGGGCGTTCGGGCTCTACCGTCCCTCTCTTCAT 1500
 DB 784 LeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHis 803
 QY 1501 CTGCGGCTTCCGCGCCACACCGGGGCGCACTCTCTTTACCGGCTCTCCCGTATGTCCT 1560
 DB 804 LeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProTyrValPro 823
 QY 1561 TCTCATCTCCGGACCGTGTGCACTTCGCTTCACTCTGCATGCTGCGATGGACACCG 1620
 DB 824 SerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 843
 RESULT 2
 ADX40768

ID ADX40768 standard; protein; 540 AA.
 AC ADX40768;
 DT 21-APR-2005 (first entry)
 DE HBV polymerase protein #11.
 XX
 XX Immune stimulation; polymerase; enzyme.
 XX Hepatitis B virus.
 OS
 XX W02005012502-A2.
 PN
 XX 10-FEB-2005.
 XX
 XX 29-MAR-2004; 2004WO-US009510.
 XX
 XX 28-MAR-2003; 2003US-0458026P.
 XX
 XX (EPIM-) EPIMUNE INC.
 XX
 XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
 PI
 XX WPI; 2005-132661/14.
 DR
 XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
 CC response comprises identifying variants of a peptide epitope 8-11 amino
 PT acids in length comprising primary anchor residues of the same HLA class
 PT I binding motif.
 XX
 XX Disclosure; Page 380-385; 458pp; English.
 PS
 XX The invention relates to a method of identifying a candidate peptide
 CC epitope which induces an HLA class I CTL response against variants of the
 CC peptide epitope, comprising identifying, from a particular antigen of an
 CC infectious agent, variants of a peptide epitope comprising primary anchor
 CC residues of the same HLA class I binding motif. The method is useful for
 CC identifying a candidate peptide epitope, which induces an HLA class I CTL
 CC response against variants of the peptide epitope. This sequence
 CC represents an HBV polymerase protein used in the scope of the invention.
 XX
 XX Sequence 540 AA;
 SQ

Alignment Scores:
 Pred. No.: 1.68e-214 Length: 540
 Score: 2810.00 Matches: 527
 Percent Similarity: 98.34% Conservative: 5
 Best Local Similarity: 97.41% Mismatches: 8
 Query Match: 47.57% Indels: 2
 DB: 9 Gaps: 0

US-10-761-006A-1 (1-3215) x ADX40768 (1-540)

QY 1 CTCGCAACATTCACCAAGCTCTCTAGATCCGAGGTGAGGCGCTATATTTTCCTGC 60
 DB 1 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 20
 QY 61 TGGTGCTCCAGTCCGGAACAGTAAACCTGTCGACTACTGCTCTCCCATATCGTC 120
 DB 21 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 40
 QY 121 AATCTTCTCGAGGACTGGGACCTCGACCGCAACATGGAGAACACAAATCAGGATTCTT 180
 DB 41 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 60
 QY 181 AGACCCCTGCTGCTGTACAGCGGGTTCCTGCTGACGAAGATCCTCACATACC 240
 DB 61 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 80
 QY 241 GCAGAGCTCTAGACTCTG-GTGACTCTCTCAATTTCTAGGGGGAGCACCCACGTGTC 299
 DB 81 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Ser 100

QY 300 CTGGCCAAATTCGCAAGTCCCAACCTCCCAATCACTCACCAACCTTGTGCTCTCAATTT 359
 DB 100 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 120
 QY 360 GTCTGGCTATCGCTGGATGTCTGGGGGTTTTTATCATATTTCTCTTTCATCTGCTGC 419
 DB 120 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 140
 QY 420 TATGCCCTCATCTTCTGTTGTTCTTCTGACTACCAAGGTATGTTGCCCTTTGCTCTC 479
 DB 140 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 160
 QY 480 TACTTCCAGGAACATCAACACCAAGCAGCGGCCCATCAAGACCTGCACGACTCTGCTGC 539
 DB 160 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 180
 QY 540 AAGGAAACTCTAGCTTTCCCTCTGTTGCTGTACAAACCTTCGACGGAACCTGCACCT 599
 DB 180 rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 200
 QY 600 GTATTCCCATCCCATCATCTCGGCTTTTCGCAAGATTCTTATGGAGTGGGCGCTCAGTCC 659
 DB 200 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 220
 QY 660 GTTCTCTGCTGCTCAGTTTACTAGTGCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCA 719
 DB 220 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 240
 QY 720 CTGTTTGGCTTTCACTATATGATGATGATGATGATGATGATGATGATGATGATGATG 779
 DB 240 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 260
 QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTG 839
 DB 260 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyLleHisLeuAsnPr 280
 QY 840 TAATAAAACCAAAAGCTGGGCTACTCCCTTAACTTCAATGGGATATGTAATGGAGTTG 899
 DB 280 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 300
 QY 900 GGGTACTTTACCGCAGGAACATATTTGTAATAAACTCAAGCAATGTTTTCGAAACTGCC 959
 DB 300 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 320
 QY 960 TGTAAATAGACTATTGATTGGAAGTATGTCAAGAAATTTGGGTCTTTTGGGCTTTTGC 1019
 DB 320 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 340
 QY 1020 TGCCCTTTTACCAATGTGCTATCTGCTCTGATGCTTTATATGATGATGATGATGATGAT 1079
 DB 340 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 360
 QY 1080 TAAGCAGCTTTTCACTTTCTCGCAACTTACAGGCTTTTCTGTGTAAACAATATCTGAA 1139
 DB 360 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 380
 QY 1140 CTTTACCCCTGGTCCCGCAACGCTCCGCTCTGCAAGTGTGCTGACGCAACCC 1199
 DB 380 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 400
 QY 1200 CACTCGATGGGCTTGGGCATAGCCATAGCCGCTTGGTGGAACTTTTCTGGCTCTCT 1259
 DB 400 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 420
 QY 1260 GCCGATCCATACCTCGGAACCTCTAGAGCTGTTGTTTGTCTCCAGCGCTCTGGAGCAA 1319
 DB 420 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLys 440
 QY 1320 ACTTATCGGAACCGCAACTCTGTTGCTCTCTCGAAATACACCTCTCTTCCATGGCT 1379
 DB 440 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 460


```
Db 340 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 360
QY 1080 TAAGCAGGCTTTCACCTTCTCCCACTTACAAAGCGCTTCTGTGTAAACATATCTGAA 1139
Db 360 rIysGlnAlaPheThrPheSerProThrTyrIysAlaPheLeuCysIysGlnTyrLeuAs 380
QY 1140 CTTTACCCCGTGTCCCGCAACGGTCCGGTCTCTCGCAAGTGTGTGTGACGCAACCCC 1199
Db 380 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 400
QY 1200 CACTCGATGGGCTTGGCCATAGGCCATCAGCGCATGTGGAACCTTTCTGGCTCTCT 1259
Db 400 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 420
QY 1260 GCCGATCATACTCGGGAACCTCTAGCAGCTTGTGTTGCTCGCAGCGGCTCGAGCAAA 1319
Db 420 uProIleHisThrAlaGluLeuAlaCysPheAlaArgSerArgSerGlyAlaIy 440
QY 1320 ACTTATCGGAACCGCAACTCTGTTGCTCTCTCGGAAATACACTCTCTTCCATGGCT 1379
Db 440 sLeulleGlyThrAspAsnSerValValLeuSerArgIysTyrThrSerPheProTrpLe 460
QY 1380 GCTAGGGTGTGTCGCAACTGGATCTCGCGGGACGTCCTTGTGTACTAGTCCCGTCGGC 1439
Db 460 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 480
QY 1440 GCTGAATCCCGGAGCGACCGCTCTCGGGCCGTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db 480 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 500
QY 1500 TCTGCGGTCGCGCGACCGACCGCGGCGCACCTCTCTTACGGGCTCCCGGTATGTGCC 1559
Db 500 sLeuProPheArgProThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 520
QY 1560 TTCTCATCTGCGGACCGGTGTCACCTTCGCTTCACTCTGCACTGCGATGAGACACC 1619
Db 520 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 540
QY 1620 G 1620
Db 540 o 540
RESULT 4
ADX40772
ID ADX40772 standard; protein; 843 AA.
XX AC ADX40772;
XX DT 21-APR-2005 (first entry)
XX DE HBV polymerase protein #15.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis B virus.
XX PN WO2005012502-A2.
XX PD 10-FEB-2005.
XX PF 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (EPIM-) EPIMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
```

```
PT I binding motif.
XX Disclosure; Page 380-385; 458pp; English.
CC The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of an
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HBV polymerase protein used in the scope of the invention.
SQ Sequence 843 AA;
Alignment Scores:
Pred. No.: 3,31e-214 Length: 843
Score: 2807.00 Matches: 528
Percent Similarity: 98.33% Conservative: 3
Best Local Similarity: 97.78% Mismatches: 8
Query Match: 47.52% Indels: 2
DB: 9 Gaps: 0
US-10-761-006A-1 (1-3215) x ADX40772 (1-843)
QY 4 CACAACTTCCCAAGCTCTGTAGATCCCAAGGTGAGGGCCCTATATTTTCTGTCTGG 63
Db 305 HisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCysTrp 324
QY 64 TGGCTCCAGTGTCCGGAACAGTAAACCTCTGTTCCGACTACTGCCTCTCCCATATCGTCAAT 123
Db 325 TrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsn 344
QY 124 CTTCTCAGAGACTGGGACCTCGACCGAAACATCGAGAACACACATCAGATTCTCTAGG 183
Db 345 LeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArg 364
QY 184 ACCCTCTCTGTTTACAGGGGGGTTTCTCGTGTGACAGAATCTCTCACAATACCCGA 243
Db 365 ThrProAlaArgValThrGlyValPheLeuValAspIysAsnProHisAsnThrThr 384
QY 244 GAGTCTAGACTCTG-GTGGACTTCTCAATTTTCTAGGGGGAGCACCCACGTTGTCCTG 302
Db 385 GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-SerTr 404
QY 303 GCCAAATTCGAGTCCCAACCTCCAACTCACTCACCACCTCTGTGCTCTCAATTTGTC 362
Db 404 pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe 424
QY 363 CTGGCTATCGGTGATGTCTGGGGGTTTATCATATTCCTCTCATCTGCTGCTAT 422
Db 424 rTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMe 444
QY 423 GCCTCATCTCTGTGTTGTTCTTGACTACCAAGGTATGTGCCCCCTTTGCTCTCTAC 482
Db 444 tProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerTh 464
QY 483 TTCAGGAGAACATCAACCAACGACGCGGGCCCATGCAAGACCTGACGACTCTCTGCTCAAG 542
Db 464 rSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSerAr 484
QY 543 GAAACTCTACGTTTCCCTCTGTTGCTGTACAAAACCTTCCGAGCGGAACCTGCATCTGTA 602
Db 484 gAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgIysLeuHisLeuTy 504
QY 603 TTCCCATCCCATCATCTCTGGGCTTTTCGCAAGATTTCTATGGAGTGGGCCCTCAGTCCCTT 662
Db 504 rSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh 524
QY 663 TCTCCTGCTCAGTTTACTAGTGCATTTGTTCAAGTGGTTCGTAGGGCTTTTCCCCCATG 722
Db 524 eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHisCys 544
```

QY 723 TTTGGCTTTCAGTTATATGATGATGTTGTTGGGGGCGAAGTCTGTACAAATCTTGA 782
Db 544 sLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLeuGI 564
QY 783 GTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCTTAA 842
Db 564 uSerLeuPheThrSerIleThrAsnPheLeuSerLeuGlyIleHisLeuAsnProAs 584
QY 843 TAAACCCAAACGTTGGGGCTACTCCCTTAACTTCATGGGATATGTAATTTGGAAGTTGGGG 902
Db 584 nLysThrLysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrpGI 604
QY 903 TACTTTACCCAGAGACATATTGTACTTAAACTCAAGCAATGTTTTCGAAAACTCGCTGT 962
Db 604 yThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuProVa 624
QY 963 AAATAGACCTATTGATTGGAAAGTATGTCAAGAAATTTGGGCTTTTGGGCTTTCTGTC 1022
Db 624 lAsnSerProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAI 644
QY 1023 CCCTTTTACAAATGTGGCTATCTCTGCTTGATGCTCTTTATATGCATGTATACAATCTAA 1082
Db 644 aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerIy 664
QY 1083 GCAGGCTTTCACCTTCTCGCAACTTACAAGGCTTTCTGTGTAAACAATATCTGAACCT 1142
Db 664 sGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsnLe 684
QY 1143 TTACCCCGTTGGCCGCGGACGGTCCGCTCTCTCCCAAGTGTCTGACGCAACCCCCAC 1202
Db 684 uTyrProValAlaArgGlnArgSerGlyLeuGlyGlnValPheAlaAspAlaThrProth 704
QY 1203 TGGATGGGCTTGGCCATAGCCCATCAGCGCATGGCTGGAACTTTCTGCTCTCTCTGCC 1262
Db 704 rGlyTyrGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuPr 724
QY 1263 GATCCATCTGCGGAACCTCTACAGCTTGTGTTTGTCTCGAGCGGTCTCGGACAAACT 1322
Db 724 oIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLysLe 744
QY 1323 TATCGGAACCGACAACTCTGTTCTCTCTCGGAAATACACCTCTTTCATGGCTGCT 1382
Db 744 uIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeuLe 764
QY 1383 AGGTGTGCTGCCAACTGGATCTCTGCGCGGACGTCCTTTGTCTACGTCCCGTCGCGCT 1442
Db 764 uGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLe 784
QY 1443 GAATCCCGGAGCAGCCGCTCTCGGGCCGTTTGGGGCTCTACGTCCTCTTCTCATCT 1502
Db 784 uAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuLe 804
QY 1503 GCGTTCGCGCCGACACCGGCGGCGACCTCTCTTTACGCGGTCTCCCGCATGTGCTTC 1562
Db 804 uProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProse 824
QY 1563 TCATCTGCCGACCGTGTGCACTTCGCTTCACCTCTGCACTGCGATGGAGACCCG 1620
Db 824 rHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaIleTrpArgProPro 843
RESULT 5
ADX40776
ID ADX40776 standard; protein; 843 AA.
XX
AC ADX40776;
XX
XX
DT 21-APR-2005 (first entry)
XX
DE HBV polymerase protein #19.
XX
KW Immune stimulation; polymerase; enzyme.
XX
XX Hepatitis B virus.
QS

XX
PN
XX
PD
XX
XX
PF
XX
PR
XX
PA
XX
PI
XX
DR
XX
PT
PT
PT
PT
XX
PS
XX
CC
CC
CC
CC
CC
CC
CC
CC
SQ
Alignment Scores:
Pred. No.: 8,98e-213 Length: 843
Score: 2789.00 Matches: 523
Percent Similarity: 98.15% Conservative: 8
Best Local Similarity: 96.67% Mismatches: 9
Query Match: 47.22% Indels: 2
DB: 9 Gaps: 0
US-10-761-006A-1 (1-3215) x ADX40776 (1-843)
QY 1 CTCACAAATTCACCAAGCTCTGTAGATCCAGGGTGTAGGGGCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGlnGlyProIlePheSerCys 323
QY 61 TGGTGGCTCCAGTTCGGAACAGTAAACCTGTTCGACTACTGCTCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTCGAGACTTGGGGACCTCGACCGAACATGGAGAACACAACTCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGACCCCTGCTCGTGTATACAGCGGGGTTTTCTCGTTGACAAGAACTCTCAATATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGACCCACGCTGTTTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisIleVal-Se 403
QY 300 CTGGCCAAAATTCGAGTCCCAACCTCCAATCACCTCACCACCTCTGTCTCTCCAAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 423
QY 360 GTCTGCTATCGCTGATGTGTCTCGGGGTTTTATCATATTCATATTCCTTCTCATCTCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAI 443
QY 420 TATGCTCATCTTCTTGTGTTGTTCTTCTGACTTACCAAGGTATGTTGCCGTTTGTCTCTC 479

W02005012502-A2.

10-FEB-2005.

29-MAR-2004; 2004WO-US009510.

28-MAR-2003; 2003US-0458026P.

(EPIM-) EPIMMUNE INC.

Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;

WPI; 2005-132661/14.

Identifying a candidate peptide epitope, which induces a HLA class I CTL response comprises identifying variants of a peptide epitope 8-11 amino acids in length comprising primary anchor residues of the same HLA class I binding motif.

Disclosure; Page 380-385; 458pp; English.

The invention relates to a method of identifying a candidate peptide epitope which induces an HLA class I CTL response against variants of the peptide epitope, comprising identifying, from a particular antigen of an infectious agent, variants of a peptide epitope comprising primary anchor residues of the same HLA class I binding motif. The method is useful for identifying a candidate peptide epitope, which induces an HLA class I CTL response against variants of the peptide epitope. This sequence represents an HBV polymerase protein used in the scope of the invention.

Sequence 843 AA;

CC used, particularly ex vivo, to stimulate CTL cells. These cells can be
 CC reintroduced into patients who have chronic or acute HBV infections or
 CC are carriers, especially in treatments to prevent conversion from acute
 CC to chronic infections. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 845 AA;

Alignment Scores: 2.19e-210 Length: 845
 Pred. NO.: 2759.00 Matches: 516
 Score: 97.23% Conservative: 10
 Percent Similarity: 95.38% Mismatches: 14
 Best Local Similarity: 46.71% Indels: 2
 Query Match: 2 Gaps: 0
 DB:

US-10-761-006A-1 (1-3215) x AAR70065 (1-845)

Qy	1	CTCCACCAACATTCACCAAGCTCTGCTAGATCCAGGGGTGAGGGCCCTATATTTTCCTGC	60
Db	306	LeuHisAsnIleProProAsnSerAlaArgSerGlnSerGluGlyProValPheSerCys	325
Qy	61	TGGTGGCTCCAGTTCGGGAAACAGTAACCCCTGTTCCGACTACTCGCTCTCCCATATGTC	120
Db	326	TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal	345
Qy	121	AATCTTCTCGAGGACTGGGACCCCTGCACCGAATCATGAGGACACACATCAGGATTCCT	180
Db	346	AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro	365
Qy	181	AGGACCCCTGCTCGTGTGTACAGGGGGGTTTTCTCTGTTGACAAAGATTCCTCACAAATACC	240
Db	366	ArgThrProAlaIaIaGValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr	385
Qy	241	GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGAGACCCACGTTTC	299
Db	386	ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrArgVal-Se	405
Qy	300	CTGGCAAAATTCGAGTCCCAACCTCAATCACTACCAACCTCTGTCTCTCCCAATTT	359
Db	405	rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe	425
Qy	360	GTCTGGCTATCTGCTGGATGTCTGCGCGTTTTTATCATATTCCTCTTCATCTCTGCTGC	419
Db	425	uSerTrpLeuSerLeuAspValSerAlaIaPheTyrHisIleProLeuHisProAlaIaI	445
Qy	420	TATGCTCATCTCTCTGTGTCTCTCTGACTCTCTCAAGGATGTTGCGCGTTGTCTCTC	479
Db	445	aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaIaArgLeuSerSe	465
Qy	480	TACTTCGAGGAACATCAACACCCAGCAGCGGGCCATGCAAGACCTGACGACTCTGCTGC	539
Db	465	rAsnSerArgIleIleAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe	485
Qy	540	AAGGAACTCTACGTTTCCCTCTCTGCTGTGTACAAAACCTTCGGACGGAACCTGCACCT	599
Db	485	rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisIle	505
Qy	600	GTATTTCCCATCCCATCATCTGGGCTTTTCGCAAGATTCCTATGGGAGTGGGCTCAGTCC	659
Db	505	uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr	525
Qy	660	GTTTCTCTGGCTCAGTTTACTAGTGCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	719
Db	525	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi	545
Qy	720	CTGTTTGGCTTTCAGTATATGATGATGTGTATTTGGGGGGAAGTCTGTACACATCT	779
Db	545	sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisIle	565
Qy	780	TCAGTCCCTTTTACTCTTATACCAATTTTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCT	839
Db	565	uGluSerLeuPheThrAlaValThrAsnPheLeuLeuLeuSerLeuGlyIleHisLeuAsnPr	585

Qy	840	TAATAAAACCAACGTTGGGGCTACTCCCTTAACTTATCGGATATGTAATGGAAGTTG	899
Db	585	oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr	605
Qy	900	GGGTACTTTACCGGAGACATATTGTTACTTAAACCAAGCAATGTTTTCGAAACTGCC	959
Db	605	pGlyThrIleProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuPr	625
Qy	960	TGTAATATAGACCTATTGATTGGAAGATGATCTCAAGAAATTTGTTGGCTTTTGGCTTTC	1019
Db	625	oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl	645
Qy	1020	TGCCCTTTTACACAATGTGCTCTATCTCTGCTGATGCTTTTATATCATGATATACATC	1079
Db	645	aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe	665
Qy	1080	TAAGCAGGCTTTCACCTTCTCGCCCACTTACAGGCTTTCTGTGTAAACAATATCTGAA	1139
Db	665	rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs	685
Qy	1140	CCTTTACCCCGTTGCGCGCAACGGTCCGGTCTCTGCAAGTGTTCCTGACGCAACCCC	1199
Db	685	nLeuTyrProValAlaIaIaArgGlnArgProGlyLeuCysGlnValPheAlaIaIaThrPr	705
Qy	1200	CACGTGATGGGGCTTGGCCATAGCCCATCAGCGCATGCGTGGAACTTTCTGGCTCTCT	1259
Db	705	oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaIaIa	725
Qy	1260	GCCGATCCATCTCGCGAATCTCTAGCAGCTTGTTCCTGCGAGCCGGTCTGGAGCAAA	1319
Db	725	uProIleHisThrAlaGluLeuAlaIaCysPheAlaArgSerArgSerGlyAlaIaIy	745
Qy	1320	ACTTATCGGAACCGACAACCTCTGTGCTCTCTCGGAAATACACCTCTTTCATGGCT	1379
Db	745	sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe	765
Qy	1380	GCTAGGTTGTCTGCCAATCGATCTCTGCGGGAGCTCTCTTGTCTACGTCCTCGGCGC	1439
Db	765	uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl	785
Qy	1440	GCTGAATCCCGCGGACGACCTCTCGGGCGGCTCTCGGGCGGCTCTACCGTCCCTCTTCA	1499
Db	785	aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuAr	805
Qy	1500	TCTCGCTTCCGGCGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1559
Db	805	gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr	825
Qy	1560	TTCTCATCTGCGGACCGGTGTGCACTTCTGCTTTCACCTCTGCACTGCGATGAGACACC	1619
Db	825	oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaIaIaTrpArgProPr	845
Qy	1620	G 1620	
Db	845	o 845	
RESULT 8			
ID	ADX40775	standard; protein; 843 AA.	
XX	ADX40775;		
AC	ADX40775;		
XX	21-APR-2005	(first entry)	
DT	21-APR-2005	(first entry)	
XX	HBV polymerase protein #18.		
DE	HBV polymerase protein #18.		
XX	Immune stimulation; polymerase; enzyme.		
KW	Hepatitis B virus.		
OS	Hepatitis B virus.		
XX	WO2005012502-A2.		
PN	10-FEB-2005.		
XX			
PD			

XX 29-MAR-2004; 2004WO-US009510.
PF
XX 28-MAR-2003; 2003US-0458026P.
PR
XX (BPIW-) EPIMUNE INC.
PA
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
PI MPI; 2005-132661/14.
XX
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
XX Disclosure; Page 380-385; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HBV polymerase protein used in the scope of the invention.
XX
XX Sequence 843 AA;
SQ

Alignment Scores:
Pred. No.: 1.14e-209 Length: 843
Score: 2750.00 Matches: 517
Percent Similarity: 97.04% Conservative: 8
Best Local Similarity: 95.56% Mismatches: 15
Query Match: 46.55% Indels: 2
DB: 9 Gaps: 0

US-10-761-006a-1 (1-3215) x ADX40775 (1-843)

QY 1 CTCACAAATCCACCAAGCTCTGTAGATCCAGGGTGAGGGGCTATATTTCTCTGC 60
DB
304 LeuHisAsnIleProSerSerAlaArgSerLysSerGluGlyProLeuPheProCys 323

QY 61 TGTGGCTCAGTTCGGGAACAGTAACCTGTTCCGACACTACTGCTCTCCCATATCGTC 120
DB 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343

QY 121 AATCTTCTCAGAGACTGGGGACCTGCACCGAACATCGGAGAACACACATCAGGATTCCT 180
DB 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlnHisAsnIleArgIlePro 363

QY 181 AGGACCCCTCTCGTGTACAGCGGGGTTTTCTCGTTGACAGAATCTCCACAATACC 240
DB 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383

QY 241 GCAGAGTCTAGACTCTG-GTGGAATTCTCTCAATTTCTAGGGGGAGACCCACGTTCTC 299
DB 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnThrGlnVal-Se 403

QY 300 CTGGCCAAATTCGAGTCCCACTCCCACTCACCACCACTCTGTCTCTCCCAATTT 359
DB 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 423

QY 360 GTCTGCTATCTCGATGTCTCGGGGTTTTATCATATTTCTTCATCTCTCTCTGC 419
DB 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaH 443

QY 420 TATGCTCATCTTCTTGTGGTCTTCTTGAGCTACCAAGGTATGTTGGCGGTTTGCTTC 479
DB 443 aMetProHisLeuLeuValGlySerGlyLeuGlnArgTyrValAlaArgLeuSerSe 463

QY 480 TACTTTCAGGAACATCAACACAGACGCGGGCCATGCAAGACTCTCAGACCTCTCTGC 539
DB

DB 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAAACTTACGTTTCCCTCTTGTGTGTGTACAAAACTTCGGAGCGAAACTGCACATT 599
DB 483 rLysHisLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTCCCATCCATCATCTCGGCTTCGCAAGATTCCTATATGGGAGTGGGCTCAGTCC 659
DB 503 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
QY 660 GTTTCCTCTCGGCTCAGTTTACTAGTCCCATTTTGTTCAGTGGTTCGTAGGCTTTCCTCCCA 719
DB 523 OpheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValCysArgAlaPheProHI 543
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTGTPATTGGGGCGAAGCTCTGTACAAATCT 779
DB 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
DB 563 uGluSerLeuPheThrSerIleThrAsnPheMetLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAAACAAACGTTGGGCTACTCCCTTAACTTTCATGGGATATGTAATTCGAAAGTTG 899
DB 583 OasnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY 900 GGGTACTTTACCGCAGGAACATATTGTACTTAAACTCAAGCAATGTTTTTCGAAAACTGCC 959
DB 603 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACCTATTGATTGGAAAGTATGTCAAAGAAATTTGGGTCTTTTGGGCTTTCG 1019
DB 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCTCTTTTACACAATGGGCTATCTCTGCTTGTATGCTCTTATATGTCATGTATACAATC 1079
DB 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAGCAGGCTTTCACCTTCTCGCCCAACTTACAAGGCTTTCGTGTGTAAACAATATCTGAA 1139
DB 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 683
QY 1140 CCTTTACCCGGTTGCCGGCAAGGTCGGTCTCTGCCAAGTGTTCGTGAGCGCAACCCC 1199
DB 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
QY 1200 CACTGGATGGGCTTGGCCATAGGCCATCAGCGCATGGGCTGGAACCTTTCTGCTCTCTCT 1259
DB 703 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 723
QY 1260 GCCGATCCATATCGCGAACTCTAGCAGCTTGTGTTTGTTCGAGCGGCTCTGGAGCAAA 1319
DB 723 uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaTh 743
QY 1320 ACTTATCGGAACCGCAACTCTGTGCTCTCTCTCGAAATACACCTCTTTCATGGCT 1379
DB 743 rLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrple 763
QY 1380 GCTAGGGTGTCTGCCCACTGGATCTCTCGCGGAGCTCTTGTGTCTAGCTCCCGCTGGC 1439
DB 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProserAl 783
QY 1440 GCTGAATCCCGCGAGCAGCCCGTCTCGGGCGGTTTGGGGCTCTACCGTCCCTTCTTCA 1499
DB 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuAr 803
QY 1500 TCTGCGGTTCCGCGCGACACCGGGCGCACCTCTCTTTACGGGCTCTCCCGCTATGTGC 1559
DB 803 gLeuSerPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
QY 1560 TTTCTCATCTCGCGGACCGGTGTGCATTTCCGCTTTCACCTCTGCACGTCCGATGAGACACC 1619
DB 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843

QY	1620 G 1620	61	TGCTGGCTCCAGTTCGGAAACAGTAAACCTGTTCGCACTACTGCTCTCCCATATGCTC	120
DB	843 o 843	21	TriPLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal	40
RESULT 9		121	AATCTTCTCGAGACTGGGGACCTCCACCGACATGGAGAACACACATCAGGATTCCT	180
ADM68212		41	AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro	60
ID	ADM68212 standard; protein; 539 AA.	181	AGGACCCCTGCTGCTGTACAGCGGGGTTTTCTCGTTGACAGAATCCTCAATPACC	240
XX	ADM68212;	61	ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr	80
XX	03-JUN-2004 (first entry)	241	GCAGAGCTAGACTCG-GRFGACTTCTCTCAATTTTCTAGGGGGAGCACCCAGCTGTC	299
XX	GRP94 protein related Hepatitis B virus protein, SEQ ID NO 4.	81	ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se	100
XX	GRP94; chaperone; DNA polymerase; Hepatitis B virus; HBV.	300	CTGGCCAAATTCGACAGTCCCAACCTCCCAATCACCACCACTCTGTCTCCCAATTT	359
XX	Hepatitis B virus.	100	rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe	120
XX	KR2002061721-A.	360	GTCTGGCTATCGCTGGATGTCTGGCGGTTTTATCATATTCTCTTCATCTCTGCTGC	419
XX	25-JUL-2002.	120	uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl	140
XX	17-JAN-2001; 2001KR-00002647.	420	TATGCTCATCTTCTTGTGTTCTCTGACTACCAAGTATGTTGCCCGTTTGTCTCTC	479
XX	17-JAN-2001; 2001KR-00002647.	140	aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe	160
XX	(MEDE-) MEDEXBIO CO.	480	TACTTCAGGAACATCAACCAGCAGCGGGCCATCGACACCTGCGACACTCTCTGCTC	539
XX	Cho YH, Kim SS, Noh HM, Shin HJ;	160	rThrSerArgAsnIleAsnHisGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe	180
XX	WPI; 2003-117827/11.	540	AAGGAACTCTACGTTTCTCTGTTGTGTACAAAACCTTTCGAGCGGAACTGCATTT	599
XX	N-PSDB; ADM68211.	180	rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe	200
XX	Grp94 acting as chaperone of DNA polymerase of hepatitis b virus and	600	GTATTCCTCATCTCTGCGCTTTCGCAAGATTCCTATGGAGTGGCGCTCAGTCC	659
XX	expression vector system for preparing a DNA polymerase in active state	200	uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyGlyGlyLeuSerPr	220
XX	of hepatitis b virus.	660	GTTTCTCTGCTGCTAGTTACTAGTGCATTTGTTTCAGTGTGTTGCTAGGCTTTCCCA	719
XX	Disclosure; SEQ ID NO 4; 15pp; Korean.	220	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi	240
XX		720	CTGTTTCTGCTTTCAGTTATATGATGTGATTTGGGGCGGAGTCTGTACAACTCT	779
XX		240	scysLeuAlaPheSerLysMetAspValValLeuGlyAlaLysSerValGlnHisLe	260
XX		780	TGAGTCTCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC	839
XX		260	uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr	280
XX		840	TAATAAACCAACGTTGGGCTACTCTCTTAACCTCATGGGATATGTAATTTGAAGTTG	899
XX		280	oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr	300
XX		900	GGGTACTTTTACCGCAGGAACATATTGTACTAAAACCTCAAGCAATGTTTTTCGAAACTGCC	959
XX		300	pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr	320
XX		960	TGTAATAGACTATTGATTGGAAAGTATGTCAAGAATTTGGGTCTTTTGGGCTTTGC	1019
XX		320	oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl	340
XX		1020	TGCCCCCTTTTACAAATGTGCTATCTGCTGCTTGTATGCTTTATATGATGTATACATC	1079
XX		340	aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe	360
XX		1080	TAAGCAGCTTTTCACTTTCTCGCAACTTACAAAGCCCTTTCTGTGTAAACAATATCTGAA	1139
XX		360	rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi	380
XX		1140	CCTTTTACCCCGTTGCCCGCAACGGTCCGGTCTCTGTGCAAGTGTGTTGTGACGCAACCCC	1199

Alignment Scores:		Length:	539
Pred. No.:	9, 9e-209	Matches:	517
Score:	2737.50	Conservative:	7
Percent Similarity:	96.86%	Mismatches:	15
Best Local Similarity:	95.56%	Indels:	3
Query Match:	46.34%	Gaps:	1
DB:	7		

US-10-761-006A-1 (1-3215) x ADM68212 (1-539)

QY	1	CTCCCAACATTTCCACAGCTCTGCTAGATCCCGGGTGGAGGGCCCTATATTTCTCGC	60
DB	1	LeuHisHisIleSerProSerProAlaArgSerGlnSerGlnGlyProIlePheSerSer	20

Db	380	sLeuTyProValAlaArg---ArgThrAlaLeuCysGlnValPheAlaAspAlaIatnPr	399
Qy	1200	CAC TGGATGGGCGCTGGCCATAGCCCATCGCGCATGGCTGGAACTTTCTGGCTCTCTCT	1259
Db	399	oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe	419
Qy	1260	GCCGATCCATACTCGGAACTCCTAGCAGCTTTGTTGCTGCGACGCGTCTGGAGCAAA	1319
Db	419	uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy	439
Qy	1320	ACTTATCGGAACCGACAACCTGTGTGCTCTCTCGGAAATACACCTCCTTCCATGGCT	1379
Db	439	sLeuIleGlyThrAspAsnSerValLeuSerArgLysTrpThrSerPheProItrPLe	459
Qy	1380	GCTAGGCTGTCTGCCAACTGGATCCTCGCGGACGTCCTTTGCTACGTCCTCGTCGCG	1439
Db	459	uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThr-Tyr-PheValTyValProSerAl	479
Qy	1440	GCTGAATCCCGCGACGACCGCTCTCGGGCCGCTTGGGGGCTCTACCGTCCTCTTCA	1499
Db	479	aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuIleArgProLeuLeuHi	499
Qy	1500	TCTGCGCTTCCGGCGGACCAACGGGGCGCACCTCTCTTTACGCGGTCTCCCGTATGTGCC	1559
Db	499	sLeuArgPheArgProThrThrGlyArgThrSerLeuTyAlaValSerProSerValPr	519
Qy	1560	TTCTCATCTGCGGACCGTGTGCACCTTCGCTTACCTCTGCACCTCGCATGGAGACCACC	1619
Db	519	oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaIatrpArgProPr	539
Qy	1620	G 1620	
Db	539	o 539	
RESULT 10			
ADX40771			
ID	ADX40771	standard; protein; 842 AA.	
AC	ADX40771;		
XX			
XX	21-APR-2005	(first entry)	
XX			
XX	HBV polymerase protein #14.		
XX			
XX	Immune stimulation; polymerase; enzyme.		
KW			
XX	Hepatitis B virus.		
OS			
XX	WO2005012502-A2.		
XX			
XX	10-FEB-2005.		
XX			
XX	29-MAR-2004; 2004WO-US009510.		
XX			
XX	28-MAR-2003; 2003US-0458026P.		
XX			
XX	(EPIM-) EPIMUNE INC.		
PA			
XX			
XX	Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;		
PI			
XX			
XX	WPI; 2005-132661/14.		
DR			

CC	infectious agent, variants of a peptide epitope comprising primary anchor				
CC	residues of the same HIA class I binding motif. The method is useful for				
CC	identifying a candidate peptide epitope, which induces an HIA class I CTL				
CC	response against variants of the peptide epitope. This sequence				
CC	represents an HBV polymerase protein used in the scope of the invention.				
XX					
SQ	Sequence 842 AA;				
Alignment Scores:					
Pred. No.:	1.95e-208	Length:	842		
Score:	2734.50	Matches:	517		
Percent Similarity:	96.86%	Conservative:	7		
Best Local Similarity:	95.56%	Mismatches:	15		
Query Match:	46.29%	Indels:	3		
DB:	9	Gaps:	1		
US-10-761-006A-1 (1-3215) x ADX40771 (1-842)					
Qy	1	CTCACAAACATTCACC	AAGCTCTGCTAGACTCCAGGTGAGGGCGCTATATTTTCTCTGC	60	
Db	304	LeuHisIleSerProAlaArgSerGlnSerGluGlyProIlePheSerSer	323		
Qy	61	TGFTGGCTCCAGTTCCGAACAGTA	AAACCTTGTTCCGACTACTGCCCTCCTCCCATCTCCTC	120	
Db	324	TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal	343		
Qy	121	AATCTTCTCAGGAGTCTGGGACCTGCACCGAACATGGAGAACACAACATCAGGATTCTCT	180		
Db	344	AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro	363		
Qy	181	AGGACCCCTGCTCGTGTTCACAGCGGGGTTTTCTCGTTGCAAGAATCTCTCACAATACC	240		
Db	364	ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr	383		
Qy	241	GCAGACTCTAGACTCTG- GTGGACTTCTCTCAATTTTCTAGGGGAGCACCACCGTGTTC	299		
Db	384	ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se	403		
Qy	300	CTGGCCAAAATTTCGACGTCCCACACCTCCAATCACTCACCAACCTTGTCTCTCAAATTT	359		
Db	403	rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnIle	423		
Qy	360	GTCTCGCTCATCTCGCTGGATGTCTCGGGCGGTTTTTATCATATTCCTCTCTCATCTGTCTGC	419		
Db	423	uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl	443		
Qy	420	TATGCTCATCTTCTGTGTGGTTCTCTCGACTACCAAGGTATGTTGCCCGTTGTCTCTC	479		
Db	443	aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe	463		
Qy	480	TACTTCAGGNACATCAACCACGACGCGGGCCATCAAGACCTGCACGACTCCTGTCTC	539		
Db	463	rThrSerArgAsnIleAsnHisGlnHisGlyThrMetGlnAspLeuHisAspSerCysE	483		
Qy	540	AAGGAAACTCTACCTTTTCCCTCTTGTGTGCTGTACAAAACCTTCGGACGGAACCTGCATTT	599		
Db	483	rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe	503		
Qy	600	GTATTCACATCCCATATCTCGGCTTTTCGAAGATTCTCTATGGAGTGGGCTCAGTCC	659		
Db	503	uTy-SerHisProIleIleLeuGlyPheArgLysIleProMetGlyGlyLeuSerPr	523		
Qy	660	GTITTCCTCGCTCAGTTTACTAGTGCATTTGTTTCAGTGCTTCGTAGGGCTTTCCCCCA	719		
Db	523	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi	543		
Qy	720	CTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGCGCAAGTCTGTACAACTCT	779		
Db	543	sCysLeuAlaPheSerTyrMetAspApValValLeuGlyAlaLysSerValGlnHisLe	563		
Qy	780	TGAGTCCCTTTTACCTATAATCAAAATTTCTTTTGTCTTTGGGTATACATTTAAACCC	839		
Db	563	uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr	583		

```
Qy 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACCTTCATGGATATGTAATTCGAAGTTG 899
Db 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
Qy 900 GGGTACTTTACCGCAGGAACATATTGACTATAAACTCAAGCAATGTTTTCGAAAACCTGCC 959
Db 603 pGlyThrLeuProGlnGluHisIleValLeuLysIleIysGlnCysPheArgLysLeuPr 623
Qy 960 TGTAAATAGACTATTGATTCGAAAGTATGTCAAAGAAATTCGGTCTTTTGGGCTTTGC 1019
Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Qy 1020 TGCCCCCTTTACAAATGTGCTACTCCGCTTCATGCTTATATGATGATGATGATGATGATG 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
Qy 1080 TAAGCAGGCTTTCACTTTCTCGCAACTTACAAGGCTTTCTGTGTAAACAATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
Qy 1140 CTTTATACCCGTTCCCGGCAACGGTCCGGTCTCTGCAAGTGTGTGCTGACGCAACCCC 1199
Db 683 sLeuTyrProValAlaArg---ArgThrAlaLeuCysGlnValPheAlaAspAlaThrPr 702
Qy 1200 CACTGGATGGGCTTGGCCATAGGCATAGGCATAGGCATAGGCATAGGCATAGGCATAGGC 1259
Db 702 oThrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 722
Qy 1260 GCCATCCATACCTCGGAACCTCTAGCAGCTTTGTTTCTCGCAGCGCTCTGGAGCAAA 1319
Db 722 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLys 742
Qy 1320 ACTTATCCGAAACCGCAACTCTGTGTCTCTCTCGGAAATACACCTCTTTTCCATGGCT 1379
Db 742 sLeuIleGlyThrAspAsnSerValAlaLeuSerArgLysTyrThrSerPheProTrpLe 762
Qy 1380 GCTAGGCTGTGTCGCAACTGATGATCTCGGGGAGCTCTTTGCTACGTCCTCGTCCG 1439
Db 762 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrTyrPheValTyrValProSerAl 782
Qy 1440 GCTGAATCCCGGACGACCGCTCTCGGGGCGTTTGGGGCTCTACCGTCCCTCTTTCA 1499
Db 782 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuIleArgProLeuLeuHi 802
Qy 1500 TCTGCGGTTCCGGCCGACACGCGGCGCACCTCTCTTTACGCGGTCTCCCGTATGTGCC 1559
Db 802 sLeuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 822
Qy 1560 TTCTCATCTGCGGACCGGTGTCATCTGCTTACCTCTGCACTGCGGATGAGACACACC 1619
Db 822 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 842
Qy 1620 G 1620
Db 842 o 842
RESULT 11
ADX40774
ID ADX40774 standard; protein; 825 AA.
XX
AC ADX40774;
XX
DT 21-APR-2005 (first entry)
XX
DE HBV polymerase protein #17.
XX
KW Immune stimulation; polymerase; enzyme.
XX
OS Hepatitis B virus.
XX
PN WO2005012502-A2.
XX
```

```
PD 10-FEB-2005.
XX
PF 29-MAR-2004; 2004WO-US009510.
XX
PR 28-MAR-2003; 2003US-0458026P.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX
XX WPI; 2005-132661/14.
XX
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.
XX
XX Disclosure; Page 380-385; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of an
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HBV polymerase protein used in the scope of the invention.
XX
XX Sequence 825 AA;
SQ
Alignment Scores:
Pred No.: 9,76e-205 Length: 825
Score: 2688.00 Matches: 504
Percent Similarity: 95.92% Conservative: 13
Best Local Similarity: 93.51% Mismatches: 21
Query Match: 45.51% Indels: 2
DB: Gaps: 0
US-10-761-006a-1 (1-3215) x ADX40774 (1-825)
Qy 7 AACATTCCCAAGCTCTGCTAGATCCAGGGTGAGGGGCTATATTTCTGCTGGTGG 66
Db 288 HisLeuSerThrThrLysArgGlnSerSerSerGlyHisAlaValGluThrCysTrpTrp 307
Qy 67 CTCAGATTCCGGAACAGTAAACCTGTTCCGACTACTGCTCTCCATATGCTCAATCTT 126
Db 308 LeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsnLeu 327
Qy 127 CTCGAGGACTGGGACCTCGACCGAACATGAGAACACATCAGGATTCCTAGGACC 186
Db 328 LeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArgThr 347
Qy 187 CTGCTCTGTTACAGGCGGGTCTTCTGTTGACAAGATCTTCACAAATACCCAGAG 246
Db 348 ProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThrThrGlu 367
Qy 247 TCTAGACTCTG-CTGGACTCTCTCAATTTCTAGGGGAGCACCACGCTGTTCTGGCC 305
Db 368 Ser***LeuValValAspPheSerGlnPheSerArgGlySerThrGlnVal-SerTrpPr 387
Qy 306 AAAATTCCAGTCCCAACCTCCAATCACTCACCACCTCTTGCTCCTCCAATTTGCTCG 365
Db 387 oLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSerTr 407
Qy 366 GCTATCGCTGGATGTTCTGGCGGCTTTTATCATATTCTCTTCTGCTGCTGCTATGCC 425
Db 407 pLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMetPr 427
Qy 426 TCATCTTCTGTTGGTCTTCTTGGGACTACCAAGGTATGTTGCCGCTTCTGCTCTACTTC 485
Db 427 oHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerThrSe 447
Qy 486 CAGGAACATCAACCAACCAACGCGGCGCATGCAAGACCTGCACGACTCCTGCTCAAGGAA 545
```



```
Db 447 rArgAsnIleAsn***GlnHisGlyThrMetGlnAspLeuHisAspSerCysSerIysHis 467
546 ACTTACTGTTCCCTCTCTGTGTGTACAAAACCTTCGGACGGAACCTGCACCTGTATTTC 605
Db 467 sLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeuTyrSe 487
QY 606 CCATCCCATCATCTCGGGCTTTTCGAAAGATTCCTATGGGAGTGGGCTCAGTCCGTTTCT 665
Db 487 rHisProIle***LeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPheLe 507
QY 666 CTGGGCTCAGTTACTAGTCCCATTTGTTTCAGTGGTTTCGTAGGGCTTCCCCACATGTTT 725
Db 507 uLeuAlaGlnPheThrSerIleAlaIleCysSerValValArgAlaPheProHisCysLe 527
QY 726 GGCTTTTCAGTTATATGATGATGTGTATTGGGGCGCAAGTCTGTACAAATCTTCGAGTC 785
Db 527 uAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGluAl 547
QY 786 CCTTTTACCTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACCTTAATAA 845
Db 547 aLeuTyrThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAsnLy 567
QY 846 AACCAACGTTGGGCTACTCCCTTAACCTTCATGGGATATGTAATTGGAAGTTGGGGTAC 905
Db 567 sThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrpGlyTh 587
QY 906 TTTTACGACATGGCTATCTCGCTGTATGCTTTTATCAAGCAATGTTTTCGAAAACGCTCTAAA 965
Db 587 rLeuProGlnAspHisIleValGlnLysLeuLysGlnCysPheArgLysLeuProValAs 607
QY 966 TAGACCTATTGATGAAAGTATGTCAAGAATTGGGTCTTTTGGGCTTTGCTGCCCC 1025
Db 607 nArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaIlePr 627
QY 1026 TTTTACACATGGCTATCTCGCTGTATGCTTTTATCAAGCAATGTTTTCGAAAACGCTCTAAA 1085
Db 627 oPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLysGl 647
QY 1086 GGCTTTACCTTTCTCGCAACTTACAAAGGCTTTCTGTGTAAACAATATCTGAACCTTTA 1145
Db 647 nAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsnLeuTy 667
QY 1146 CCCCCTTCCCGCAACGGTCCGGTCTCTGCCAAGTGTCTGACGACCAACCCCACTGG 1205
Db 667 rProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProThrGl 687
QY 1206 ATGGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACCTTTCTGCTCCTCTGCGGAT 1265
Db 687 YTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuProIl 707
QY 1266 CCATCTCGCGAACTCCTAGCAGCTTGTGTTGCTCGACGCGGTCTGGAGCAAACTTAT 1325
Db 707 eHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLysLeuIl 727
QY 1326 CGGAACCGCAACTCTGTGTCTCTCTCGAAATATACCTCTCTTCATGCTGCTAGG 1385
Db 727 eGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpTrpLeuGl 747
QY 1386 GTCTGCTGCCAATCTGGATCTCGCGGAGCTCTTTGCTAGTCTCCGTCGCGCTGAA 1445
Db 747 yCysAlaIleAsnTrpLeuLeuArgGlyThrSerPheValTyrValProSerAlaLeuAs 767
QY 1446 TCCCGGCGACGACCCGCTCTCGGGCGGTTTGGGGCTCTACCGTCCCTCTTCTCATCTGCC 1505
Db 767 nProAlaAspAspProSerArgLysLeuGlyLeuTyrArgProLeuLeuHisLeuPr 787
QY 1506 GTTCCGGCGACACGCGGGGCACTCTCTTTAGCGGCTCTCCCGTATGTGCTTCTCTCA 1565
Db 787 oPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSerHi 807
QY 1566 TCTGCGGACCGGTGTCACCTTCCTTCACCTCTGACGTCGATGGAGACCCG 1620
```

```
Db 807 sLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 825
RESULT 12
ADX40773
ID ADX40773 standard; protein; 827 AA.
XX
AC ADX40773;
XX
DT 21-APR-2005 (first entry)
XX
DE HBV polymerase protein #16.
XX
KW Immune stimulation; polymerase; enzyme.
XX
OS Hepatitis B virus.
XX
PN WO2005012502-A2.
XX
PD 10-FEB-2005.
XX
PF 29-MAR-2004; 2004WO-US009510.
XX
PR 28-MAR-2003; 2003US-0458026P.
XX
PA (EPIM-) EPIMUNE INC.
XX
PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
WPI; 2005-132661/14.
XX
PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
PS Disclosure; Page 380-385; 458pp; English.
XX
CC The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variant of the peptide epitope. This sequence
CC represents an HBV polymerase protein used in the scope of the invention.
SQ Sequence 827 AA;
Alignment Scores:
Pred. No.: 5.57e-204 Length: 827
Score: 2678.50 Matches: 504
Percent Similarity: 95.91% Conservative: 12
Best Local Similarity: 93.68% Mismatches: 18
Query Match: 45.34% Indels: 5
DB: Gaps: 1
US-10-761-006A-1 (1-3215) x ADX40773 (1-827)
QY 19 AGCTCTGCTAGATCCACGGGTGAGGG-----CCTATATTTTCTGCTGGTGGCTC 69
Db 291 ThrSerLysArgGlnSerSerSerGlyHisAlaValGluLeuHisProCysTrpTrpLeu 310
QY 70 CAGTTCCGGAACAGTAAACCTGTTCGACTACTGCTCTCCCATATCGTCAATCTTCTC 129
Db 311 GlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsnLeuLeu 330
QY 130 GAGGACTGGGGCCCTGACCGAATGAGACACAACTCAGGATTCCTAGGACCCCT 189
Db 331 GluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArgThrPro 350
QY 190 GCTCGTGTTCACAGGGGGGTTTCTCGTTGACAAAGATCCTCACAATACCGCAGAGTCT 249
Db 351 AlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThrGluSer 370
```


QY 250 AGACTCTG-GTGGACTCTCTCAATTTCTAGGGGAGCACCACGCTGTTCTCTGGCCAAA 308
DB |||||
QY 371 ArgLeuValValAspPheSerGlnPheSerArgGlySerThr**Val-SerTrpProLy 390
DB |||||
QY 309 ATTCGCAGTCCCAACTCCAACTCACACCACTCTGTCTCTCCAAATTTGCTCTGGCT 368
DB |||||
QY 390 sPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSerTrpLe 410
DB |||||
QY 369 ATCCCTCGATGTGTCTCGGGGTTTTATCATATTCCTCTTCATCTCTGCTATGCTCA 428
DB |||||
QY 410 uSerLeuAspValSerAlaAlaPheTyHieProLeuHisProAlaAlaMetProHi 430
DB |||||
QY 429 TCTCTTGTGTTCTCTGACCTACCAAGGTATGTTGCCGTTGCTCTCTACTTCCAG 488
DB |||||
QY 430 sLeuLeuValGlySerSerGlyLeuProArgTyValAlaArgLeuSerSerThrSerAr 450
DB |||||
QY 489 GAACATCAACACACGACGAGGGGCATGCAAGACCTGCAGACTCTCTGCTCAAGGAACT 548
DB |||||
QY 450 gAsn***AsnTyHieGlnHisGlyThrMetGlnAspLeuHisGluSerCysSerArgAsnLe 470
DB |||||
QY 549 CTACGTTTCCCTCTGTGTGTGTACAAACCTTCGAGCGGAACTGCATCTGTATTCCTCA 608
DB |||||
QY 470 uTyHieValSerLeuLeuLeuTyHieLysThrPheGlyArgGlyLeuHisLeuTyHieSerHi 490
DB |||||
QY 609 TCCCATCATCTCGGCTTCGCAAGATCTCTATGGAGTGGGCTCAGTCCGTTCTCCT 668
DB |||||
QY 490 sProileileuGlyPheArgGlyLysIleProMetGlyValGlyLeuSerProPheLeuLe 510
DB |||||
QY 669 GGCTCAGTTTACTAGTGCATTTGTTCTAGTGTGTTCTGAGGCTTTCCCACTGTTGGC 728
DB |||||
QY 510 uAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHisCysLeuAl 530
DB |||||
QY 729 TTTTCACTTATATGATGATGTGTATTTGGGGCGAAGTCTGTACACATCTTGAGTCCCT 788
DB |||||
QY 530 aPheSerTyMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGluSerLe 550
DB |||||
QY 789 TTTTACCTCTATTACCAATTTCTTGTCTTGTGGGTATACATTTAAACCTTAATAAC 848
DB |||||
QY 550 uPheThrAlaIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProHisLysfth 570
DB |||||
QY 849 CAACGTTGGGGTACTCCCTTAACCTCATCGGATATGTAATTAATGGAAGTTGGGTACTTT 908
DB |||||
QY 570 rLysArgTrpGlyTyHieSerLeuAsnPheMetGlyTyHieValIleGlySerTrpGlyThrLe 590
DB |||||
QY 909 ACCGAGGAACATATTTGACTAAACCTCAAGCAATGTTTTCGAAACCTGCTGTAATAG 968
DB |||||
QY 590 uProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuProValAsnAr 610
DB |||||
QY 969 ACCTATTGATTGGAAAGTATGTCAGAAATTTGGGTCTTTGGGCTTTGGTCCCTTT 1028
DB |||||
QY 610 gProileAspTrpLysValCysGln**IleValGlyLeuLeuGlyPheAlaAlaProPh 630
DB |||||
QY 1029 TACCAATGTGCTATCTGCTGCTGATGCTTTATATCATGTATACAACTCTAAGCAGGC 1088
DB |||||
QY 630 eThrGlnCysGlyTyHieProAlaLeuMetProLeuTyHieAlaCysIleGlnAlaLysGlnAl 650
DB |||||
QY 1089 TTTTCACTTCTCGCCAACTTCAAGGCTTTCTGTGTAAACAATATCTGAACCTTTACCC 1148
DB |||||
QY 650 aPheThrPheSerProThrTyHieAlaPheLeuCysLysGlnTyHieLeuHisLeuTyHiePr 670
DB |||||
QY 1149 GCTTCCCGGAGAACGGTCCGCTCTCTGCAAGTGTCTGACGCAACCCCTGAGT 1208
DB |||||
QY 670 oValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProThrGlyTr 690
DB |||||
QY 1209 GGGCTTGGCCATAGCCATCAGCGATCGCTGGAACCTTTCTGGCTCTCTGCGCATCCA 1268
DB |||||
QY 690 pGlyLeuAlaIleGlyHisArgGlyMetArgGlyThrPheValAlaProLeuProIleHi 710
DB |||||
QY 1269 TACTCGGAACCTCTAGCAGCTTGTGTTGCTCGCAGCGGCTGTGAGCAAACTTATCGG 1328
DB |||||
QY 710 eThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLysLeuIleG1 730
DB |||||

QY 1329 AACCGAACACTGTGTGTCTCTCTCGAAATACACCTCTCTCCATGGCTGCTAGGGTG 1388
DB |||||
QY 730 yThrAspAsnSerValValLeuSerArgLysTyHieThrSerPheProTrpLeuLeuGlyCy 750
DB |||||
QY 1389 TGTGTCAACTGATCTCTCGCGGAGCGTCTCTTTGTCTAGTCCCGTCCGCGCTGAATCC 1448
DB |||||
QY 750 sAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyHieValProSerAlaLeuAsnPr 770
DB |||||
QY 1449 CGCGGAGCAGCCGCTCTCGGGCGGCTTTGGGGCTCTACCGTCCCTCTCTCATCTGCCGCT 1508
DB |||||
QY 770 oAlaAspAspProSerArgGlyArgLeuGlyLeuTyHieArgProLeuLeuHisLeuProPh 790
DB |||||
QY 1509 CCGGCCGACACACCGGCGCACCTCTCTTTACGCGGTCTCCCGTATGTGCTTCTCACT 1568
DB |||||
QY 790 eArgProThrThrGlyArgThrSerLeuTyHieAlaValSerProSerValProSerHisLe 810
DB |||||
QY 1569 GCGGAGCCGCTGTGCTCTTCCCTTCCCTCTGACGCTCGCATGGAGACCACCG 1620
DB |||||
QY 810 uProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 827
DB |||||
RESULT 13
AAE04708
ID AAE04708 standard; protein; 842 AA.
XX AC
XX AAE04708;
XX
XX 04-SEP-2001 (first entry)
XX
XX Hepatitis B virus FRI strain genotype G HBpol protein.
XX
XX HBV genotype G; precore; HBpol; polymerase; envelope protein; preS1;
KW preS2; surface antigen; HBsAg; HBx protein; vaccine; liver disease;
KW hepatitis; liver cancer; HBcAg; core antigen.
XX
XX Hepatitis B virus.
OS
XX
XX WO200138498-A2.
XX
XX 31-MAY-2001.
XX
XX 21-NOV-2000; 2000WO-US032108.
XX
XX 24-NOV-1999; 99US-0167206P.
XX
XX (PHAR-) PHARMASSET INC.
XX (INNO-) INNOGENETICS NV.
XX
XX Stuyver L, Schinazi R, De Gendt S, Van Geyt C, Zoulim F, Fried M,
PI Robsaw R;
XX
XX WPI; 2001-367676/38.
XX N-PSDB; AAD09091.
XX
XX Novel hepatitis B virus genotype G, nucleic acids encoding virus,
PT polypeptides encoded by nucleic acids, useful for preparing vaccine to
PT treat or prevent the hepatitis B virus genotype G infection in a subject.
XX
XX Example; Fig 6; 84pp; English.
XX
XX The present invention relates to hepatitis B virus (HBV) strain FRI,
CC genotype G DNA encoding PreCore/Core protein, HBpol, envelope (preS1,
CC preS2 and surface antigen HBsAg) and HBx proteins. HBV genotype G nucleic
CC acids and polypeptides are useful for diagnosing, prognosing and treating
CC infections caused by HBV genotype G. They can be used in a vaccine to
CC treat or prevent HBV genotype G infection. The HBV genotype G derived
CC nucleic acids and antibodies are useful for detecting HBV genotype G in a
CC sample or diagnosis of HBV genotype G infection. The presence of HBV
CC genotype G statistically correlates with the presence of liver damage
CC and/or liver cancer in the subject. The HBV genotype G core insert
CC peptide encoding nucleic acid is useful for designing monitoring assays
CC to study and predict the evolution of anti-HBe and anti-HBc antibodies
CC and HBsAg (genotype G e antigen) in patients infected with HBV. The
CC antibodies or antigens of HBV genotype G are useful for identifying a

CC stage of liver disease caused by HBV genotype G. The present sequence is
 CC hepatitis B virus (HBV) strain FRI, genotype G polymerase protein, HBpol
 XX
 SQ Sequence 842 AA;

Alignment Scores:
 Pred. No.: 1,066-203 Length: 842
 Score: 2675.00 Matches: 496
 Percent Similarity: 96.30% Conservative: 25
 Best Local Similarity: 91.68% Mismatches: 19
 Query Match: 45.29% Indels: 2
 DB: 4 Gaps: 0

US-10-761-006A-1 (1-3215) x AAE04708 (1-842)

QY	1	CTCCACAACTCCACCAAGCTCTGCTAGATCCAGGCTGAGGGGCTATATTTCTCTGC	60
DB	303	LeuTyr:SerileProSerSerThrLysSerGlnSerGlnGlyProValPheSerCys	322
QY	61	TGTTGGCTCCAGTTCGGGAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTC	120
DB	323	TrpTrpLeuGlnPheArgAspSerGluProCysSerAspTyrCysLeuSerHisLeuVal	342
QY	121	AATCTTCTCAGGACTGGGGACCTCGACCGAACATGGAGAACACACATCAGGATTCCT	180
DB	343	AsnLeuLeuGlnAspTrpGlyProCysThrGluHisGlyGluHisIleArgIlePro	362
QY	181	AGGACCCCTGCTCGTGTACAGCGGGGTTTTCTCGTTGACAAAGAAATCCTCAATAACC	240
DB	363	ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr	382
QY	241	GCAGAGTCTAGACTCTG-GTGGACTTCTCAATTCCTCAACCACTCTGTGCTCCCAATTT	299
DB	383	AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerAlaArgVal-Se	402
QY	300	CTGGCCAAATTCGAGTCCCAACTCCCAATCACTCAACCACTCTGTGCTCCCAATTT	359
DB	402	rTTPProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe	422
QY	360	GTCTGGCTATCGCTGGATGTGTCTGGGGTTTTATCATATTCCTCTTCATCTCTCTGC	419
DB	422	uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaH	442
QY	420	TATGCTCATCTTCTGTGGTCTTCTGACTACCAAGTATGTTGGCGTTGTGCTTC	479
DB	442	aMetProHisLeuLeuValGlySerGlyLeuSerArgTyrValAlaArgLeuSerSe	462
QY	480	TACTTCCAGGAACATCAACACAGCAGCGGGCCATGCAAGACCTGCACGACTCCTGCTC	539
DB	462	rAspSerArgIleLeuAspHisGlnTyrGlyThrLeuGlnAsnLeuHisAspSerCysSe	482
QY	540	AAGAAACTCTAGCTTTCCCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	599
DB	482	rArgGlnLeuTyrValSerLeuMetLeuLeuTyrLysThrPheGlyArgLysLeuHisLe	502
QY	600	GTATTCCTCATCATCTCTGGCTTTCGCAAGATTCCTATGGAGTGGGCTCTAGTCC	659
DB	502	uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr	522
QY	660	GTCTTCTCTGGCTCAGTTTACTAGTGCATTTCTGCTAGTGTCTGTTAGGCTTTCCCCCA	719
DB	522	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProH	542
QY	720	CTGTTTGGCTTTTCAGTTATATGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	779
DB	542	sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe	562
QY	780	TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCT	839
DB	562	uGluSerLeuTyrThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr	582
QY	840	TAATAAAACCAACGTTGGGGCTACTCCCTTAACCTTCACTGATATGTAATTCGGAAGTTG	899

DB	582	oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValleGlySerTr	602
QY	900	GGGTACTTTTACCGCAGGAACATATTGTACTAAACTCAAGCAATGTTTTCGAAATCGCC	959
DB	602	pGlyThrLeuProGlnGluHisIleThrGlnLysIleLysGlnCysPheArgLysLeuPr	622
QY	960	TGTAATATAGACCTATTGATTGGAAAGTATGTCAAGAATTGGGTCTTTTGGCTTTGC	1019
DB	622	oValAsnArgProIleAspTrpLysValCysGlnArgIleThrGlyLeuLeuGlyPheAl	642
QY	1020	TGCCCCCTTTTACACAATGTGGCTATCCTCGCTTGATGCTCTTATATGATATATCAATC	1079
DB	642	alaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAl	662
QY	1080	TAAGCAGGCTTTCACCTTCTCGCCAACTTCAAGGCTTCTGTGTGTGTGTGTGTGTGTGT	1139
DB	662	alysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrMetAs	682
QY	1140	CCTTTACCCCGTTGCCCGCAACGGTCCCGTCTCTGCCAAGTGTGTGCTGACGCAACCCC	1199
DB	682	nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr	702
QY	1200	CACGTGATGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACCTTCTGCTGCTCTCT	1259
DB	702	oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe	722
QY	1260	GCGATTCATATCGCGAACTCTCTAGCAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1319
DB	722	uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys	742
QY	1320	ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGGAATACACCTCTTTCATCGGCT	1379
DB	742	sLeuIleGlyThrAspAsnSerValValleuSerArgLysTyrThrSerPheProIlePr	762
QY	1380	GCTAGGCTGTGTGCGCAACTGATCTCTGCGCGGAGCTCTTGTCTAGCTCCCGTCGGC	1439
DB	762	uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl	782
QY	1440	GCTGAATCCCGCGGACGACCCGCTCTCGGGGCGGTTTGGGGCTCTACCGTCCCTTCTTCA	1499
DB	782	leuAsnProAlaAspProSerArgGlyArgLeuGlyLeuCysArgProLeuLeuAr	802
QY	1500	TCTGCGTTCGCGCGGACCGACCGCGGCGACCTCTCTTTACGGGCTCTCCCGTATGTGC	1559
DB	802	gluProPheLeuProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr	822
QY	1560	TTCTCATCTGCGCGGACCGGTGTGCATCTTCCCTTCCACCTCTGCACGTGCGATCGAGACC	1619
DB	822	oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValThrTrpLysProPr	842
QY	1620	G 1620	
DB	842	O 842	
RESULT 14			
ID	AAG66918 standard; protein; 842 AA.		
XX	AAG66918;		
AC	19-OCT-2001 (first entry)		
DT	HBV pres1/32/HBS region protein.		
XX	Hepatitis B virus; HBV; preCore; Core; pres1; pres2; HBS; HBX; HBPol;		
KW	HBSAg; antiviral; vaccine; genotype G; genotyping; HBcAg; HBeAg.		
XX	Hepatitis B virus.		
XX	W0200140279-A2.		
XX	07-JUN-2001.		

PF 20-NOV-2000; 2000WO-EP011526.
 XX
 PR 03-DEC-1999; 99EP-00870252.
 PR 07-DEC-1999; 99US-0169287P.
 XX
 PA (INNO-) INNOGENETICS NV.
 XX
 PI Stuyver L, Van Geyt C, De Gendt S;
 XX
 DR WPI; 2001-374785/39.
 DR N-PSDB; AAH77562.
 XX
 PT Novel isolated and/or purified hepatitis B virus polypeptide and
 PT polynucleotide sequences that are phylogenetically different from HBV
 PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
 PT therapy.
 XX
 PS Claim 16; Fig 1; 94pp; English.
 XX
 CC The invention relates to the complete nucleic acid sequence of a new
 CC human hepatitis B virus (HBV) genotype, provisionally named genotype G.
 CC This genotype was found with a high prevalence in patients chronically
 CC infected with HBV and residing in Europe and the USA. The invention
 CC relates to a fully defined sequence of 3248 nucleotides as given in
 CC specification, a sequence with 92% identity to the given sequence, or
 CC sequence that is degenerate to the mentioned sequences. These
 CC polynucleotides are useful for HBV genotyping. The proteins encoded by
 CC the polynucleotides are useful for detecting antibodies in a biological
 CC sample. Ligands that bind to the proteins and antibodies directed against
 CC the proteins are useful for detecting the proteins and for detecting
 CC HBsAg and HBsAg (precursor proteins). They are also useful for
 CC preparing a vaccine or medicament for treating HBV infections. The
 CC present sequence is encoded by the genome of HBV genotype G strain FRI
 XX
 SQ Sequence 842 AA;

Alignment Scores:
 Pred. No.: 1,066-203 Length: 842
 Score: 2675.00 Matches: 496
 Percent Similarity: 96.30% Conservative: 25
 Best Local Similarity: 91.68% Mismatches: 19
 Query Match: 45.23% Indels: 2
 DB: 4 Gaps: 0

US-10-761-006a-1 (1-3215) x AAG66918 (1-842)

QY 1 CTCCACAACTTCCACCAAGCTCTCTAGATCCCAAGGCTGAGGGCCCTATATTTTCCTGC 60
 DB 303 LeuTyrSerIleProProSerSerThrIysSerGlnSerGlnGlyProValPheSerCys 322
 QY 61 TGGTGGCTCCAGTTCGGAAACAGTAACCCCTGTTCCGACTACTGCTCTCCCATATCGTC 120
 DB 323 TrpTrpLeuGlnPheArgAspSerGluProCysSerAspTyrCysLeuSerHisLeuVal 342
 QY 121 AATCTTCTCGAGGACTGGGACCTGACCGAACATGAGAACACAACTCAGGATTCCT 180
 DB 343 AsnLeuLeuGlnAspTrpGlyProCysThrGluHisGlyGluHisIleArgIlePro 362
 QY 181 AGGACCTCTGCTGCTGCTACAGCGGGGTTTCTCTGTTGACAAAGATCCTCACAAATACC 240
 DB 363 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 382
 QY 241 GCAGAGTCTAGACTCTG-GTGACTTCTCAATTTCTAGGGGAGCACCCACGCTGTC 299
 DB 383 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerAlaArgVal-Se 402
 QY 300 CTGGCCAAATTCGAGTCCCAACTCCCAATCACTCACCACCTCTGTGCTCCCAATTT 359
 DB 402 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 422
 QY 360 GTCCTGGCTATCGCTGGATGTCTGCGGCGGTTTATCATATTTCTCTTCATCTGCTGC 419
 DB 422 userTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 442

QY 420 TATGCCTCATCTTCTTGTGGTTCTTCTGACTACCAAGGTATGTTGCCGTTGTGCTCTC 479
 DB 442 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 462
 QY 480 TACTTCAGAAACATCAACACAGCAGCGGGCCATGCAAGACCTGCAGACCTGCCTGCTC 539
 DB 462 rAspSerArgIleLeuAspHisGlnTyrGlyThrLeuGlnAsnLeuHisAspSerCysSe 482
 QY 540 AAGGAAACTCTAGCTTTCCTCTTGTGCTGACAAACCTTCGGACGGAAACTGCACCT 599
 DB 482 rArgGlnLeuTyrValSerLeuMetLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 502
 QY 600 GTATTCCTCCATCCATCATCTCGGCTTTCGCAAGATTCCTATGGAGTGGGCTCATGTC 659
 DB 502 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 522
 QY 660 GTTTCCTCGGCTCAGTTACTAGTGCATTTGTCAGTGGTTCGTAGGCGCTTCCCA 719
 DB 522 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 542
 QY 720 CTGTTTGGCTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGATG 779
 DB 542 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 562
 QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTCTTCTTGGGTATACATTTAAACCC 839
 DB 562 uGluSerLeuTyrThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 582
 QY 840 TAATAAACAACAGTGGGCTACTCCCTTAACCTTACCTCATGGATATGATGATGATGATG 899
 DB 582 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 602
 QY 900 GGGTACTTTACCGCAGGACATATTGACTAAACCTCAAGCAATGTTTTCGAAACCTGCC 959
 DB 602 pGlyThrLeuProGlnGluHisIleThrGlnLysIleLysGlnCysPheArgLysLeuPr 622
 QY 960 TGTAATAGACCTATTGATTGGAAGTATGTCAAAGAAATTGTGGTCTTTTGGGCTTTGC 1019
 DB 622 oValAsnArgProIleAspTrpLysValCysGlnArgIleThrGlyLeuLeuGlyPheAl 642
 QY 1020 TGCCCCCTTTTACAAATGTCGCTATCCCTGCTGATCCCTTATATGATGATGATGATGATC 1079
 DB 642 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAl 662
 QY 1080 TAAGCAGCTTTCATCTTCTCGCAACTTCAAGGCTTTCTGTGTAACAATATCTGAA 1139
 DB 662 aLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrMetAs 682
 QY 1140 CTTTACCTCCCTTCCCGCAACCGTCCGGTCTCTGCAAGTGTTCGTCAGCAACCC 1199
 DB 682 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 702
 QY 1200 CACTGGATGGGCTTGGCCATAGCCCATCAGCCGATGCTGGAACCTTTCGGTCTCTCT 1259
 DB 702 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 722
 QY 1260 GCCATCCATCTACCGGAACCTCTAGCAGCTTGTGTCGCGAGCGGCTCGAGAGCAA 1319
 DB 722 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 742
 QY 1320 ACTTATCGAAACCGCAACTCTGTTGCTCTCTCGGAAATACACCTCTTTTCCATGCT 1379
 DB 742 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTiple 762
 QY 1380 GCTAGGCTGCTGCTCCAACTCGATCTCGCGGGAGCTCTTGTCTACGTCCTCGCTCGGC 1439
 DB 762 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 782
 QY 1440 GCTGATCCCGGAGCAGCCGCTCTCGGGGCGTGTGGGCTCTACCGCTCCCTCTTCTCA 1499
 DB 782 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuCysArgProLeuLeuAr 802

QY 1500 TCTGCGTTCCGGCCGACCGGGGCGACCTCTCTTTAGCGGTCTCCCGTATGTGCC 1559
 Db 802 gLeuProPheLeuProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 822
 QY 1560 TTCTCATCTCCGGACCGTGTGCACTTTCACCTCTGACGTCGCGATGGAGACACC 1619
 Db 822 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValThrTrpLysProPr 842
 QY 1620 G 1620
 Db 842 o 842
 RESULT 15
 ADX40761
 ID ADX40761 standard; protein; 832 AA.
 XX
 AC ADX40761;
 XX
 DT 21-APR-2005 (first entry)
 XX HBV polymerase protein #4.
 DE
 XX Immune stimulation; polymerase; enzyme.
 KW
 XX Hepatitis B virus.
 OS
 XX WO2005012502-A2.
 PN
 XX 10-FEB-2005.
 PD
 XX 29-MAR-2004; 2004WO-US009510.
 PF
 XX 28-MAR-2003; 2003US-0458026P.
 PR
 XX (EPIM-) EPIMUNE INC.
 PA
 XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
 PI WPI; 2005-132661/14.
 DR
 XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
 PT response comprises identifying variants of a peptide epitope 8-11 amino
 PT acids in length comprising primary anchor residues of the same HLA class
 PT I binding motif.
 XX
 PS Disclosure; Page 380-385; 459pp; English.
 XX
 CC The invention relates to a method of identifying a candidate peptide
 CC epitope which induces an HLA class I CTL response against variants of the
 CC peptide epitope, comprising identifying, from a particular antigen of an
 CC infectious agent, variants of a peptide epitope comprising primary anchor
 CC residues of the same HLA class I binding motif. The method is useful for
 CC identifying a candidate peptide epitope, which induces an HLA class I CTL
 CC response against variants of the peptide epitope. This sequence
 CC represents an HBV polymerase protein used in the scope of the invention.
 XX
 SQ Sequence 832 AA;
 Alignment Scores:
 Pred. No.: 4.97e-202 Length: 832
 Score: 2654.00 Matches: 492
 Percent Similarity: 95.56% Conservative: 25
 Best Local Similarity: 90.94% Mismatches: 23
 Query Match: 44.93% Indels: 2
 DB: 9 Gaps: 0
 US-10-761-006a-1 (1-3215) x ADX40761 (1-832)
 QY 1 CTCACAACTTCCAAAGCTCTGTAGATCCAGGTCAGGGGCTATATTTCTCTGTC 60
 Db 293 LeuHisAsnProProAsnSerAlaArgSerGlnGlyGluArgProValPheProCys 312
 QY 61 TGTGGGTCCTCCGACAGTAAACCCCTGTTCGGACTACTCTCCCATATCGTC 120

Db 313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisLeVal 332
 QY 121 AATCTTCTCGAGACTGGGGACCCCTGCACCGAACATGGAGAACACAACTCAGGATTCCT 180
 Db 333 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisLeArgIlePro 352
 QY 181 AGGACCCCTGCTCGTGTTCACAGCGGGGTTTTCTCGTTTGACAAGAATCCTCACATACC 240
 Db 353 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 372
 QY 241 GCAGAGCTTAGACTCTG-CTGGACTCTCTCAATTTCTAGGGGAGACCCACGCTGTC 299
 Db 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnHisArgVal-Se 392
 QY 300 CTGGCCAAATTCGAGTCCCAACCTCCAATCACCCTCACCACCTCTGCTCTCCCAATTT 359
 Db 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 412
 QY 360 GTCTGGCTATCGCTGGATGTGTCTGGCGGTTTTATCATATATTCCTTCATCTCTGCTGC 419
 Db 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaSe 432
 QY 420 TATGCTCATCTCTTGTGTTGTTCTCTCGACTTACCAGGTATGTTGCCGTTTGTCTC 479
 Db 432 rMetProHisLeuLeuValGlySerThrGlyLeuSerArgTyrValAlaArgValSerSe 452
 QY 480 TACTTCCAGGAACATCAACCCAGCAGCGGGCCATGCAAGACCTGCACGACTCTCTGCTC 539
 Db 452 rAsnSerArgIlePheAsnHisGlnArgGlyThrMetGlnAsnLeuHisAspTyrCysSe 472
 QY 540 AAGGAACTCTAGTTCCTCTTGTGCTGTGTACAAAACCTTCGGACGGAACCTGACATT 599
 Db 472 rArgAsnLeuTyrValSerLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe 492
 QY 600 GTATTCCTCATCTCCTCGGCTTCGCAAGATTCTATGGGAGTGGGCTCAGTCC 659
 Db 492 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512
 QY 660 GTTTCCTCTGGCTCAGTTTACTAGTGCCTATTTGTTTCAGTGGTTTCGTAGGCTTTCC 719
 Db 512 OpheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 532
 QY 720 CTGTTTGGCTTCAGTTATATGATGATGTGTATTTGGGGGGAAGTCTGTACAACTCT 779
 Db 532 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 552
 QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTAAACCC 839
 Db 552 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 572
 QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACCTTCATGGGATATGTAATTTGGAAGTTG 899
 Db 572 OasnLysThrLysArgTyrGlyTyrSerLeuHisPheMetGlyTyrValIleGlySery 592
 QY 900 GGGTACTTTACGCGAGGAACATATGTACTATAAACCAAGCAATGTTTTCGAAAACTGCC 959
 Db 592 rGlySerLeuProGlnAspHisIleIleGlnLysIleLysGluCysPheArgLysLeuPr 612
 QY 960 TGTAAATAGACCTATTGATTGGAAAGTATGTCAAAGAAATTTGGGGCTTTTGGGCTTTGC 1019
 Db 612 oIleAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
 QY 1020 TGCCTCTTTTACAAATGTGGCTATCTCTGCTGTATGCTCTTATATGATGATATACAATC 1079
 Db 632 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 652
 QY 1080 TAAGCAGGCTTTCACHTTCTCGCAACTTCAAGGCTTTCTGTGTAAACAATATCTGAA 1139
 Db 652 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 672
 QY 1140 CCTTTACCCCGTTCGCCGCAACGGTCTCTGCCAAGTGTGTTGCTGACCAACCC 1199

```

Db      672 nLeuTyrProValalaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
Qy      1200 CACTGGATGGGCTTGGCCATAGGCCATAGCGCATGGCTGGAAACCTTTCTGGCTCCTCT 1259
Db      692 oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaProLe 712
Qy      1260 GCCGATCCATACTCGGAACCTCCTAGCAGCTTGTGTTGCTCGCAGCCGGTCTGGAGCAAA 1319
Db      712 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732
Qy      1320 ACTTATCGGAACCGACAACCTCTGTGCTCTCTCGGAATACACCTCCTTTCCATGGCT 1379
Db      732 nIleLeuGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 752
Qy      1380 GCTAGGGTGTCTGCCAACTGATCCTCGCGGGACGTCCTTTCTACGTCCTCGGTCGGC 1439
Db      752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 772
Qy      1440 GCTGAATCCCGCGGACGACCGCTCTCGGGGCGGTTTGGGGCTCTACCGTCCCTCTTCA 1499
Db      772 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyIlePheArgProLeuLeuAr 792
Qy      1500 TCTGCCGTTCGGCGCGACACCGCGCGCACCTCTCTTTACCGGCTCTCCCGTATGTGCC 1559
Db      792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr 812
Qy      1560 TTCTCATCTGCCGACCGTGTGCACTTCGCTTTCACCTCTGCACGTCGGATGGAGACCACC 1619
Db      812 oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832
Qy      1620 G 1620
Db      832 o 832

```

Search completed: December 1, 2005, 23:35:15
Job time : 520.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 1, 2005, 23:15:45 ; Search time 106 Seconds
(without alignments)
5836.545 Million cell updates/sec

Title: US-10-761-006A-1

Perfect score: 5907

Sequence: 1 CTCACACATTCACCAAG.....CCTCAGGCCACGAGTGGAA 3215

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlpl
-Q=/cgn2.1/USPTO_spool_p/US10761006/runat_01122005_113953_26882/app_query.fasta_1.3399
-DB=PIR -QFMT=FASTA -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=EXPT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10761006@cgn 1 1 163 @runat_01122005_113953_26882 -NCPUs=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2807	47.5	843	1 JDVLVR	DNA-directed DNA p
2	2789	47.2	843	2 S35527	DNA-directed DNA p
3	2750	46.6	843	1 S43491	DNA-directed DNA p
4	2734.5	46.3	842	1 JDVLVS	DNA-directed DNA p
5	2688	45.5	825	2 T13473	DNA-directed DNA p
6	2678.5	45.3	827	2 T13468	DNA-directed DNA p
7	2654	44.9	832	2 S71785	DNA-directed DNA p
8	2652	44.9	832	1 JDVLVB	DNA-directed DNA p
9	2651	44.9	832	1 S47406	DNA-directed DNA p
10	2643	44.7	832	1 JDVLVA	DNA-directed DNA p
11	2643	44.7	843	1 JDVLJ3	DNA-directed DNA p
12	2642	44.7	832	1 JDVLJ1	DNA-directed DNA p
13	2637	44.6	832	1 S20757	DNA-directed DNA p
14	2628	44.5	832	2 S67505	DNA-directed DNA p

15	2628	44.5	843	1 JDVLJ1	DNA-directed DNA p
16	2623	44.4	832	1 JDVLCP	DNA-directed DNA p
17	2610	44.2	843	1 JDVLJ2	DNA-directed DNA p
18	2608	44.2	845	1 JDVLKS	DNA-directed DNA p
19	2607	44.1	832	1 S20752	DNA-directed DNA p
20	2607	44.1	845	1 JDVLVD	DNA-directed DNA p
21	2582	43.7	843	1 JQ2229	DNA-directed DNA p
22	2230	37.8	750	1 JDVLVH	DNA-directed DNA p
23	1792	30.3	359	2 S04570	gene X/C fusion pr
24	1719	29.1	884	1 JDVL7	DNA-directed DNA p
25	1716	29.1	884	1 JDVL59	DNA-directed DNA p
26	1714	29.0	556	1 JDVL64	DNA-directed DNA p
27	1708	28.9	884	1 JDVLW8	DNA-directed DNA p
28	1698	28.7	879	1 JDVLC	DNA-directed DNA p
29	1668	28.2	881	1 JDVLS	DNA-directed DNA p
30	1649	27.9	883	1 JDVLC2	DNA-directed DNA p
31	1472	24.9	400	1 SAVLA	large surface anti
32	1452	24.6	400	2 S35528	surface antigen -
33	1391.5	23.6	445	2 S43492	large surface anti
34	1380	23.4	389	1 SAVLCP	large surface anti
35	1376	23.3	389	1 SAVLBH	large surface anti
36	1372	23.2	389	2 S67506	surface antigen -
37	1369	23.2	389	2 S20745	major surface anti
38	1368	23.2	400	1 JQ1575	large surface anti
39	1367	23.1	389	2 S32202	large surface anti
40	1363	23.1	389	2 S47407	surface antigen -
41	1363	23.1	389	2 S20753	surface antigen -
42	1361	23.0	400	1 SAVLVD	large surface anti
43	1356	23.0	389	2 S20749	surface antigen -
44	1355	22.9	384	2 T13469	large surface anti
45	1354	22.9	389	1 SAVLAH	large surface anti

ALIGNMENTS

RESULT 1

JDVLVR
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adr)
C:Species: hepatitis B virus, HBV
C>Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 09-Jul-2004
C:Accession: A00704
R:Ono, Y.; Onda, H.; Sasada, R.; Igarashi, K.; Sugino, Y.; Nishitoka, K.
Nucleic Acids Res. 11, 1747-1757, 1983
A:Title: The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype δ
A:Reference number: A93460; MUID:83168919; PMID:6300776
A:Molecule type: DNA
A:Residues: 1-843 <ONO>
A:Cross-references: UNIPROT:P03157; UNIPARC:UPI00001297FE
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:
Pred. No.: 1.13e-199 Length: 843
Score: 2807.00 Matches: 528
Percent Similarity: 98.33% Conservative: 3
Best Local Similarity: 97.78% Mismatches: 8
Query Match: 47.52% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1 (1-3215) x JDVLVR (1-843)

Qy	4	CACACATTCACCAAGCTCTGTAGATCCAGGTCAGGGCCCTATATTTCTCTGCTGG	63
Db	305	HieAsnIleProProSerSerAlaArgSerGlnserGluGlyProIlePheSerCysIrp	324
Qy	64	TGGCTCCAGTTCCGGAACAGTAAACCCCTGTTCCGACTACTGCTCTCCCATATCGTCAAT	123
Db	325	TrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsn	344
Qy	124	CTTCTCGAGGACTGGGACACCTGCACCGAATCATGAGAACACATCATGAGATTCCTAGG	183
Db	345	LeuLeuGluAspTrpGlyProCysThrGluHisGluHisAsnIleArgIleProArg	364

```
QY 184 ACCCTGCTCGTTACAGCGGGTTTTTCTCGTTGACAAAGAACTCCTCAATACCGCA 243
Db 365 ThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThrThr 384
QY 244 GAGTCTAGACTCG-GTGGACTTCTCTCAATTTCTAGGGGGAGCACCCACGTTGCCCG 302
Db 385 GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-SerTr 404
QY 303 GCCAAATTCGCGAGTCCCAACCTCCAATCACTCACCAACTCTGTCTCTCCAAATTGTG 362
Db 404 pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe 424
QY 363 CTGGCTATCCCTCGATGTGTCGGGGTTTTATCATATTCCTCTTCATCCTGCTGCTAT 422
Db 424 rTrpLeuSerLeuAspValSerAlaAlaPheTyHisIleProLeuHisProAlaAlaMe 444
QY 423 GCCTCATCTTTGTTGTTCTCTGACTACCAAGGTATGTTGCCCGTTGTGCTCTAC 482
Db 444 cProHisLeuLeuValGlySerSerGlyLeuProAr-gTyValAlaArgLeuSerSerTh 464
QY 483 TTCAGGAACATCAACACAGCACGGGCCCATGCAAGACCTGCACGACTCTCTGCTCAAG 542
Db 464 rSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSerAr 484
QY 543 GAAACTCTAGCTTCCCTCTGTTGTTGTTGTACAAAACCTTCGGACGGAACTGCACTGTGA 602
Db 484 gAsnLeuTyTrValSerLeuLeuLeuLeuTyTrLysThrPheGlyArgLysLeuHisLeuTy 504
QY 603 TTCCCATCCCATCATCTCTGGGCTTCGCAAGATTCCTATGGGAGTGGGCTCAGTCCGTT 662
Db 504 rSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh 524
QY 663 TCTCCTGGCTCAGTTACTAGTCCCTTTGTTTCAGTGGTTTCGTAGGCGTTTCCCCACTG 722
Db 524 eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCy 544
QY 723 TTTGGCTTTCAGTTATATGATGATGTGTTATTTGGGGGCGAAGCTGTCAACATCTTGA 782
Db 544 sLeuAlaPheSerTyMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGl 564
QY 783 GTCCCTTTTACCTATTTACCAATTTCTTTGTTCTTTGGGTATACATTTAAACCTTAA 842
Db 564 uSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAs 584
QY 843 TAAACCAACACGTTGGGGCTACTCCCTTAACCTTCATCGGATATGTAATTTGGAAGTTGGGG 902
Db 584 nLysThrLysArgTrpGlyTyTrSerLeuAsnPheMetGlyTyTrValIleGlySerTrpGl 604
QY 903 TACTTTACCGCAGAACATATTGTATAAAGCTCAAGCAATGTTTTCGAAAACCTGCTGT 962
Db 604 yThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuProVa 624
QY 963 AAATAGACCTATTGATTGGAAATGTCAAGAATTGGGTCCTTTGGGCTTTGCTGCTGC 1022
Db 624 lAsnSerProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl 644
QY 1023 CCCTTTTACAAATATGGCTATCCTGCTTGATGCGCTTTATATGTCATGTATACAATCTAA 1082
Db 644 aProPheThrGlnCysGlyTyTrProAlaLeuMetProLeuTyTrAlaCysIleGlnSerLy 664
QY 1083 CGAGGCTTTCACCTTCTCGCCAACTTACAAAGGCTTTCTGTGTAAACAATATCTGAACCT 1142
Db 664 sGlnAlaPheThrPheSerProThrTyTrLysAlaPheLeuCysLysGlnTyTrLeuAsnLe 684
QY 1143 TTACCCGTTGCGCGCAACGTCCTGCTCTCCCAAGTGTGTTGTCGACGACACCCCAAC 1202
Db 684 uTyTrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProTh 704
QY 1203 TGGATGGGCTTGGCCATAGGCCATCAGCGCATGGCTTGGAAACCTTTCTGCTCCTCTGCC 1262
Db 704 xGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuPr 724
```

```
QY 1263 GATCCATACTCGGAACTCCTAGCAGCTGTGTTTCTCGACGCGGTCTGGAGCAAAACT 1322
Db 724 oIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLysLe 744
QY 1323 TATCGGAACCGCAACTCTGTGTCTCTCTCGGAAATACACTCTCTTTCCATGGCTGCT 1382
Db 744 uIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLeuLe 764
QY 1383 AGGGTGTGTCGCAACTGGATCCTCGCGGAGCGTCTTTGTCTAGTCCCGTCGCGCGCT 1442
Db 764 uGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyTrValProSerAlaLe 784
QY 1443 GAATCCCGCGGACGACCGCTCTCGGGCGGTTTGGGCTCTACCGTCCCGCTTTCATCT 1502
Db 784 uAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyArgProLeuLeuLeu 804
QY 1503 GCGGTTCCCGCGCACACCGGCGGCACTCTCTTACGCGGCTCTCCCGTATGTGCTTTC 1562
Db 804 uProPheArgProThrThrGlyArgThrSerLeuTyTrAlaValSerProSerValProSe 824
QY 1563 TCATCTGCGGACCGGTGTCATTCGCTTTCACCTCTGCACGTGCGATGAGACCAACG 1620
Db 824 rHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 843
RESULT 2
S35527
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adr)
C:Species: hepatitis B virus, HBV
A:Variety: subtype adr
C:Date: 09-Dec-1993 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004
C:Accession: S35527
R:Mukaide, M.; Kumazawa, T.; Hoshi, A.; Kawaguchi, R.; Hikiji, K.
Nucleic Acids Res. 20, 6105, 1992
A:Title: The complete nucleotide sequence of hepatitis B virus, subtype adr (SRADR) and f
A:Reference number: S35527; MUID:93096607; PMID:1461746
A:Accession: S35527
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-843 <NUK>
A:Cross-references: UNIPROT:081107; UNIPARC:UPI000000BED31; EMBL:D12980; NID:g221500; PID:
A:Experimental source: subtype adr
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1992
C:Genetics:
A:Introns: 345/3
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase
Alignment Scores:
Pred. No.: 2,44e-198 Length: 843
Score: 2789.00 Matches: 523
Percent Similarity: 98.15% Conservative: 8
Best Local Similarity: 96.67% Mismatches: 9
Query Match: 47.22% Indels: 2
DB: 2 Gaps: 0
US-10-761-006A-1 (1-3215) x S35527 (1-843)
```

```
QY 1 CTCCCAACATTCACCAAGCTCTGCTAGATCCAGGGTGAGGGCGCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
QY 61 TGTGTGCTCCAGTTCGGAACAGTAAACCTGTGTCGACTACTGCTCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyTrCysLeuThrHisIleVal 343
QY 121 ATCTTCTCGAGGACTGGGGACCTGCACCGAACAATGGAGAACACAAACATCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTGTTGTACAGCGGGGTTTTTCTGTTGACAAGAAATCCTCACATAACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
```



```
QY 241 GCAGAGTCTAGACTCTG-GTGGACTCTCTCAATTTTCTAGGGGAGCACCACCGTGTTC 299
Db |||
QY 384 ThrGluSerArgLeuValAlaPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY 300 CTGGCCAAAATTCGACGTCCCAACCTCCAACTCACTCACCACCTCTGTCTCCCAATTT 359
Db |||
QY 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY 360 GTCTGCTGATCGTGATGCTGGCGGCTTTTATCATATTCCTCTCATCTCTGCTGC 419
Db |||
QY 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
QY 420 TATCCCTCATCTCTGTGTGCTCTCTGGACTACCAAGGTATGTCGCCGTTTGTCTCTC 479
Db |||
QY 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaargLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACCAACGACGCGGCCCATCGAAGCTGCACGACTCTCTGCTC 539
Db |||
QY 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAAACTCTACGTTTCCCTCTCTGTTGCTGTGTACAAACCTTCGAGCGGAACCTGCACTT 599
Db |||
QY 483 rArgAsnLeuTyrValSerLeuLeuValTyrLysThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTCCCATCCCATCTCCGCGCTTCGCAAGATTCTCATGGAGTGGCGCTCAGTCC 659
Db |||
QY 503 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
QY 660 GTTCTCTCTGCTCAGTTTACTAGTGCATTTGTTTCAGTGTGTTCTGAGGCTTTCCCCCA 719
Db |||
QY 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValargAlaPheProHi 543
QY 720 CTGTTTGGCTTTTCAGTTATATGATGATGTGGTATGTGGGCGCAAGCTGTGACAACTCT 779
Db |||
QY 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGCTTTGGGTATACATTTAAACCC 839
Db |||
QY 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAACAACCAAGCTTGGGCTACTCCCTTAACCTTCATGGGATATGTAATTTGGAAGTTG 899
Db |||
QY 593 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY 900 GGGTACTTTACCGCAGGAACATATTGTACTAAACTCAAGCAATGTTTTGGAACCTGCC 959
Db |||
QY 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACTATTGATTGGAAGTATGTCAAAGATTTGGGTCTTTTGGGCTTTCG 1019
Db |||
QY 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCTTTTACCAATGTGCTATCTCCCTGCTGATGCTTTATATGATGATGTATACAAATC 1079
Db |||
QY 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAAGCAGGCTTTTCACTTTCTCGCCAACTTACAAGGCTTTTCTGTGTAACAATATCTGAA 1139
Db |||
QY 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysGlnGlnTyrLeuHi 683
QY 1140 CTTTATCCCGTTCGCCGCAACCGTTCGGTCTCTGCAAGTTTGTGTCGCGCAACCCC 1199
Db |||
QY 683 sLeuTyrProValAlaAlaArgGlnArgSerGlyLeuCysGlnValPheGlyAspAlaThrPr 703
QY 1200 CACTGGATGGGCTGGCCATAGGCATCAGCGCATGGTGGAACTTTTCTGGCTCCTCT 1259
Db |||
QY 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgargMetArgGlyThrPheValAlaProLe 723
QY 1260 GCCGATCCATCTCGGGAACTCCTAGCAGCTTTGTTTGTCTCGACCGGTCTGGAGCAAA 1319
Db |||
QY 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 743
QY 1320 ACTTATCGGAACCGCAAACTCTGTTGCTCTCTCGGAATAACACTCTCTTCCATGGCT 1379
```

```
Db |||
QY 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
QY 1380 GCTAGAGTGTCTGCGCAACTGGATCTGCGCGGAGCGTCTTGTCTACGTCCTCGGTCCGC 1439
Db |||
QY 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
QY 1440 GCTGAATCCCGCGGACGACCCGCTCTCGGGCCGTTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db |||
QY 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
QY 1500 TCTGCGGTTCCGGCCGACACGCGGGCGACCTCTCTTACGCGGTCTCCCGTATGNGCC 1559
Db |||
QY 803 sLeuProPheGlnProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
QY 1560 TTTCTATCTCCCGGACGCGTGTGCTTCCCTTCCCTCTCAGCTCGCATCGGAGACCAACC 1619
Db |||
QY 823 oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
QY 1620 G 1620
Db |||
QY 843 O 843
Db |||
RESULT 3
S43491
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adr)
C:Species: hepatitis B virus, HBV
A:Variety: subtype adr
C:Date: 07-Sep-1994 #sequence_revision 22-Nov-1996 #text_change 09-Jul-2004
C:Accession: S43491
R:Loncarevic, I.F.; Zentgraf, H.; Schroeder, C.H.
Nucleic Acids Res. 18, 4940, 1990
A:Title: Sequence of a replication competent hepatitis B virus genome with a prex open tñ
A:Reference number: S12598; MUID:90370503; PMID:2395664
A:Accession: S43491
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-843 <LON>
A:Cross-references: UNIPROT:Q67952; UNIPARC:UPI00000EE7B3; EMBL:X52939; NID:G457780; PDB:
A:Experimental source: subtype adr
C:Genetics:
A:Gene: P; pol
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase
Alignment Scores:
Pred. No.: 1.89e-195 Length: 843
Score: 2750.00 Matches: 517
Percent Similarity: 97.04% Conservative: 8
Best Local Similarity: 95.56% Mismatches: 15
Query Match: 46.55% Indels: 2
DB: 1 Gaps: 0
US-10-761-006A-1 (1-3215) x S43491 (1-843)
QY 1 CTCCACAACTTCCACCAAGCTCTAGATCCAGGCTAGGCGCTATATTTTCCTGC 60
Db |||
QY 304 LeuHisAsnIleProProSerSerAlaArgSerLysSerGluGlyProLeuPheProCys 323
QY 61 TGTGTGGCTTCCAGTTCGCGAAACAGTAAACCCCTGTTCGCACTACTGCTCTCCCATATCCTC 120
Db |||
QY 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTCGAGGACTGGGACCCCTGCACCGAATGGAGAACACAAATCAGGATTCCT 180
Db |||
QY 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTCGTGTGTACAGCGGGGTTTTTCTCGTTGACAAATCTCCACAAATACC 240
Db |||
QY 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGTCTAGACTCTG-GTGGACTCTCTCAATTTTCTAGGGGAGCACCACCGTGTTC 299
```

Db 384 ThrGluSerArgLeuValValAaspPheSerGlnPheSerArgGlyAsnThrGlnVal-Se 403
 Qy 300 CTGGCCAAATTCGAGTCCCAACCTCCATCACCACCTCTTGTCTCCAAATTT 359
 Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 423
 Qy 360 GTCTCGCTATCGCTCGATGTGTCGCGGCTTTTATCATATTCCTTCCTCATCTGCTCC 419
 Db 423 uSerTrpLeuSerLeuAaspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
 Qy 420 TATGCTCATCTCTTCTTGTGTTCTTCTGAGCTACCAAGGTATGTTGCCGTTGTCTTC 479
 Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuGlnArgTyrValAlaArgLeuSerSe 463
 Qy 480 TACTTCCAGGAACATCAACACACGACGCGGGCCATGCAAGACCTGCACGACTCTGCTC 539
 Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
 Qy 540 AAGGAAACTCTACGTTTCCCTCTGTTGCTGTACAAAACCTTCGGACGGAAACCTGCAC 599
 Db 483 rLysHisLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
 Qy 600 GTATTCCCATCCCATCATCTCGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
 Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
 Qy 660 GTTCTCTCGGCTCAGTTACTAGTGCCATTTGTTCAAGTGGTTCGTAGGGCTTTCCCA 719
 Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValCysArgAlaPheProHi 543
 Qy 720 CTGTTTGGCTTCAGTATATGATGATGATGATGATGATGATGATGATGATGATGATG 779
 Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
 Qy 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
 Db 563 uGluSerLeuPheThrSerIleThrAsnPheMetLeuSerLeuGlyIleHisLeuAsnPr 583
 Qy 840 TAATAAACAACAAAGTTGGGGCTACCTCCCTTAACTTCATGGGATATGTAATTGGAAGTTG 899
 Db 583 oAsnLysThrLysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
 Qy 900 GGGTACTTTACCGACAGACATTTGCTATAAATCAAGCAATGTTTTCGAAACTGCC 959
 Db 603 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 623
 Qy 960 TGTAAATAGACCTATTGATTGGAAGTATGTCAAAGATTGTTGGGTCTTTTGGGCTTTGC 1019
 Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
 Qy 1020 TGCCCTTTTACAAATGTGGCTATCCTGCCTTGATGCTTATATGATGATGATGATGATG 1079
 Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
 Qy 1080 TAAGAGGCTTTCACCTTTCGCGCAACTTACAGGCTTTCGTGTAAACAATATCTGAA 1139
 Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 683
 Qy 1140 CTTTTACCGTTGCGCGCAACGGTCCGCTCTCGCAAGTGTGTTGCTGAGCGCAACCCC 1199
 Db 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
 Qy 1200 CACTGGATGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACTTCTTGCTCTCTCT 1259
 Db 703 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 723
 Qy 1260 GCGGATCCATACGCGGAACCTCTACGAGCTTGTGTTGCTCGGACGGCTCTGGAGCAA 1319
 Db 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaTh 743
 Qy 1320 ACTTATCGGAACGCAACTCTGTTGCTCTCTCGGAAATACACCTCTCTTCCATGGCT 1379
 Db 743 rLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763

Qy 1380 GCTAGGTGTGCTGCGCAACTCGATCCTCGCGGGAGCTCTTGTCTAGTCCCGTCGCG 1439
 Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
 Qy 1440 GCTGAATCCCGCGGACGACCCCGTCTCGGGCGCGTGTGGGGCTCTACCGTCCCGCTTCTTCA 1499
 Db 783 aLeuAsnProAlaAaspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuAr 803
 Qy 1500 TCTGCGCTTCCGGCGACACACGCGGCGCACCTCTCTTTACGCGGCTCTCCCGTATGTGCC 1559
 Db 803 gLeuSerPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
 Qy 1560 TTCTCATCTCGCGGACCGTGTGCACCTTCGCTTCACCTCTGACGCTCGCATGGAGCCACC 1619
 Db 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
 Qy 1620 G 1620
 Db 843 o 843
 RESULT 4
 JDVLVS
 DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adr, mutant)
 C:Species: hepatitis B virus, HBV
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
 C:Accession: S04568
 R:Rho, H.M.; Kim, K.; Hyun, S.W.; Kim, Y.S.
 Nucleic Acids Res. 17, 2124, 1989
 A:Title: The nucleotide sequence and reading frames of a mutant hepatitis B virus subtype
 A:Reference number: S04568; MUID:89183619; PMID:2928116
 A:Accession: S04568
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-842 <RHO>
 A:Cross-references: UNIPROT:P31870; UNIPARC:UPI0001297PB; EMBL:X14193
 C:Superfamily: hepatitis virus DNA-directed DNA polymerase
 C:Keywords: DNA biosynthesis; nucleotidyltransferase
 Alignment Scores:
 Pred. No.: 2,65e-194 Length: 842
 Score: 2734.50 Matches: 517
 Percent Similarity: 96.86% Conservative: 7
 Best Local Similarity: 95.56% Mismatches: 15
 Query Match: 46.29% Indels: 3
 DB: 1 Gaps: 1
 US-10-761-006A-1 (1-3215) x JDVLVS (1-842)
 Qy 1 CTCGCACACATTCACACAGCTCTCTAGATCCAGGGTGAGGGCGCTATATTTTCCTGC 60
 Db 304 LeuHisHisIleSerProSerProAlaArgSerGlnSerGlnGlyProIlePheSerSer 323
 Qy 61 TCGTGGCTCCGAGTCCGGACAGTAAACCTGCTCCGACTCTGCTCTCCCATATCGTC 120
 Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
 Qy 121 AATCTTCTCGAGACTGCGGACCTCTGCACCGAACATGGAGAACACACATCAGGATTCCT 180
 Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
 Qy 181 AGGACCCCTGCTGTTTACAGCGGGGTTTTTCTCGTTGACCAAGATCCTCACAATACC 240
 Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
 Qy 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGAGGACCCACGCTGTTTC 299
 Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
 Qy 300 CTGGGCAAAATTCGAGTCCCGCAACCTCCAATCACTCAACACCTCTTGTCTCTCCAATTT 359
 Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423

```

QY 360 GTCTGCTATCGTGTGATGCTGCGGCGTTTATCATATTCCTCTTTCATCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
QY 420 TATCCCTCATCTCTCTGTTGCTTCTCTGAGTACCAAGGTATGTTCCCGTTTGTCTTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
QY 480 TACTTCAGGAACATCAACACACAGCACGGGCCCATGCAAGACCTGCAGACCTCTGCTC 539
Db 463 rThrSerArgAsnIleAsnHisGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAAACTCTACGTTTCCCTCTGTTGCTGTCTACAAACCTTCGGACGGAAACTGCACCT 599
Db 483 rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTCCCATCCCATCTCTGGGCTTTTCGCAAGATTCCTATGGAGTGGGCGCTTCAGTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyGlyLeuSerPr 523
QY 660 GTTTCCTCTGCTCAGTTTACTAGTCCATTTGTTTCAGTGTTCGTAGGCGTTTCCGCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 839
Db 563 uGlnSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAAACCAACCTGGGCTACTCCCTTAACCTCAATGATGATGATGATGATGATGATG 899
Db 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY 900 GGGTACTTTACCGAGNACATATGTACTAAACTCAAGCAATGTTTTCGAAACTGCC 959
Db 603 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAATAAGACTATTGATTCGAAGTATGTCAAAGATTTGGGTCTTTTGGGCTTTGC 1019
Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCCCTTTTACAAATGTGCTATCTCGCTTGTATGCTTATATGATGATGATGATGAT 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAGCAGGCTTTTCACTTTCTCGCNACTTACAGGCTTTTCTGTGTAACATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
QY 1140 CTTTTACCCCGTTGCCCGCAACGGTCCGGTCTCTGCCAAGTGTGTTGTCGACGCAACCCC 1199
Db 683 sLeuTyrProValAlaArg--ArgThrAlaLeuCysGlnValPheAlaAspAlaThrPr 702
QY 1200 CACTGATGGGCTTGGCCATAGGCCATAGGCCATAGGCCATAGGCCATAGGCCATAGGCC 1259
Db 702 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 722
QY 1260 GCCGATCATACTGCGGAACCTCTAGCAGCTTTGTTGCTCGCAGCGGCTCGGAGCAA 1319
Db 722 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLy 742
QY 1320 ACTTATCGAACCGCAACTCTGTTGCTCTCTCTCGAAATACACCTCTTTCATGCTGCT 1379
Db 742 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 762
QY 1380 GCTAGGGTGTGTCGAACCTGGATCCTCGCGGGACGCTCTTGTGTCTACGTCCTCGGC 1439
Db 762 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrTyrPheValTyrValProSerAl 782
QY 1440 GCTGAATCCCGGAGGACCGCTCTCGGGGCGGTTTGGGGCTCTACCGCTCCCTCTTCA 1499

```

```

Db 782 aLeuAsnProAlaAspProSerArgGlyArgLeuGlyLeuIleArgProLeuLeuHi 802
QY 1500 TCTGCGCTTCCGGCCACACCGGGCCACCTCTCTTTAGCGGTCTCCCGGTATGCGC 1559
Db 802 sLeuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 822
QY 1560 TTTCTATCTCCGGACCGTGTGACCTTCCGCTTCCAGCTCTCAGCTGCATGGAGACAC 1619
Db 822 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 842
QY 1620 G 1620
Db 842 O 842

RESULT 5
T13473
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (isolate 29Y11HCC)
C:Species: hepatitis B virus, HBV
A:Variety: isolate 29Y11HCC
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T13473
R/Takahashi, K.; Akahane, Y.; Hino, K.; Ohta, Y.; Mishiro, S.
Arch. Virol. 143, 2313-2326, 1998
A:Title: Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinoma
A:Reference number: Z17684; MUID:99129050; PMID:9930189
A:Accession: T13473
A:Status: preliminary; translated from GB/EMBL/DBDJB
A:Molecule type: DNA
A:Residues: 1-825 <TAK>
A:Cross-references: UNIPROT:Q9YZS3; UNIPARC:UPI00000F4BF5; EMBL:AB014388; MID:93582381;
A:Experimental source: Japanese patient with hepatocellular carcinoma isolate 29Y11HCC
C:Genetics:
A:Gene: P
A:Introns: 303/3
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:
Pred. No.: 7,33e-191 Length: 825
Score: 2688.00 Matches: 504
Percent Similarity: 95.92% Conservative: 13
Best Local Similarity: 93.51% Mismatches: 21
Query Match: 45.51% Indels: 2
DB: Gaps: 0

US-10-761-006A-1 (1-3215) x T13473 (1-825)
QY 7 AACATTCCACCAAGCTCTGCTAGATCCAGGGTGGGGCGCTATATTTTCTGCTGTGG 66
Db 288 HisLeuSerThrThrLysArgGlnSerSerSerGlyHisAlaValGluThrCysTrpTrp 307
QY 67 CTCAGTTCGGAAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTCAATCTT 126
Db 308 LeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsnLeu 327
QY 127 CTCAGAGACTGGGACCGCTCCGACATCGAAGACACACATCAGGATTCCTAGGAC 186
Db 328 LeuGlnAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArgThr 347
QY 187 CTCTCTGTTTACAGCGGGGTTTTCTCTCTGACCAAGATTCCTCAATACCGCAGAG 246
Db 348 ProIleArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThrGlu 367
QY 247 TCTAGACTCTG-GTGGACTCTCTCAATTTTCTAGGGGAGCACCACCGTGTCTGCGCC 305
Db 368 Ser**LeuValAlaAspPheSerGlnPheSerArgGlySerThrGlnVal-SerTrpPr 387
QY 306 AAAATTGCGAGTCCCAACCTCCATCACTCAACCACTCTGCTCCTCCCAATTTGCTCG 365
Db 387 oLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSerTr 407
QY 366 GCTATCGTGTGATGTCTGCGGGCGGTTTTTATCATATTCCTCTTTCATCTGCTGTATGCC 425

```

||||| pleuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMetPr 427
QY 426 TCATCTCTTGTGGTCTCTCGACTACCAAGGTATGTTGCCGGTTTCTCTCTACTTC 485
Db 427 oHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerThrSe 447
QY 486 CAGGAACATCAACCACGACGCGGCCCATGCAGACTTCGACGACTCTGCTCAAGGAA 545
Db 447 rArgAsnIleAsn**GlnHieGlyThrMetGlnAspLeuHisAepSerCysSerIyshi 467
QY 546 ACTCTACGTTCCCTCTGTGTGTGTACAAACCTTCGCGAGCAATCTGCACATCTTGATTC 605
Db 467 sleuTyrValSerLeuLeuLeuLeuTyrHisThrPheGlyArgIyLeuHisLeuTyrSe 487
QY 606 CCATCCCATCATCTCGGGCTTCGCAAGATTCCTATGGAGTGGGCTCAGTCCGGTTCT 665
Db 487 rHisProIle**LeuGlyPheArgIyIleProMetGlyValGlyLeuSerProPheLe 507
QY 666 CTGGCTCAGTTTACGTAGTCGCATTTGTTTCAGTGGTTCGTAGGCTTTCCTCCACGTGTTT 725
Db 507 uLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHisCysLe 527
QY 726 GGCCTTTCAGTTATATGATGATGTGTATTTGGGGCGAAGTCTGTACAACATCTTCAGTC 785
Db 527 uAlaPheSerTyrMetAspAspValValLeuGlyAlaIySerValGlnHisLeuGluAl 547
QY 786 CCTTTTACCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCTTAATAA 845
Db 547 aLeuTyrThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAsnly 567
QY 846 AACCAACGTTGGGGCTACTCCCTTAACCTTCATGGGATATGTAATTGGAGTGGGGTAC 905
Db 567 sThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrpGlyTh 587
QY 906 TTTTACCGCAGGAACATATTGTACTAAACCTCAAGCAATGTTTTCGAAACCTGCTGTA 965
Db 587 rLeuProGlnAspHisIleValGlnLysLeuLysGlnCysPheArgLysLeuProValAs 607
QY 966 TAGACCTATTGATTGAAAGTATGTCAAGAAATTTGGGTCTTTTGGGCTTTGCTGCC 1025
Db 607 nArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAlaPr 627
QY 1026 TTTTACCAATGGGCTATCCCTGCTTATGCTTATATGATGATGATATCAATCTAAGCA 1085
Db 627 oPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLysG 647
QY 1086 GGCCTTTCACCTTCTCGCCAACTTACAAAGCCTTCTGTGTAAACAATATCTGAACCTTTA 1145
Db 647 nAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsnLeuTy 567
QY 1146 CCCCGTTGCCCGCAACGGTCCGGTCTCTGCCAAGTGTGTTGTGACGCAACCCCACTGG 1205
Db 667 rProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProThrG 687
QY 1206 ATGGGCTTGGCCATAGGCCCATCGGCATGGCTGGAACTTTCGCTCCTCTGCGCAT 1265
Db 687 yTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuProIl 707
QY 1266 CCATCTCGCGAACTCTAGCAGCTTGTGTTGCTCGCAGCCGGTCTGGAGCAAACTTAT 1325
Db 707 eHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLysLeuIl 727
QY 1326 CGGAACCGCAACTCTGTGTCTCTCTCGGAATAACACTCTCTTCCATGGCTGTAGG 1385
Db 727 eGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeuLeuG 747
QY 1386 GTGTGTCGCAACTGATCTCGCGGGGAGCTCTTTGCTAGCTCCGCTCGCGGCTGAA 1445
Db 747 yCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLeuAs 767
QY 1446 TCCCGGGGACGACCCGCTCTCGGGGGCGTGTGGGGCTCTACCGCTCCCTTCTCATCTGCG 1505

Db 767 nProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHisLeuPr 787
QY 1506 GTTCCGGCGCACCAACGGGGCGACCTCTCTTTACCGGGTCTCCCGTATGTCCTTCTCA 1565
Db 787 oPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSerHi 807
QY 1566 TCTGTCGGACCGTGTGCACCTTCGCTTCCACCTCTGCACGTTCGCATCGACACACCG 1620
Db 807 sleuProAspArgValHisPheAlaSerProLeuHisValAlaIleTrpArgProPro 825
RESULT 6
T13468
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (isolate 09D09HCC)
C:Species: hepatitis B virus, HBV
A:Variety: isolate 09D09HCC
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T13468
R:Takahashi, K.; Akahane, Y.; Hino, K.; Ohta, Y.; Mishiro, S.
Arch. Virol. 143, 2313-2326, 1998
A:Title: Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinoma
A:Reference number: Z17684; MUID:99129050; PMID:9930189
A:Accession: T13468
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-827 <TAK>
A:Cross-references: UNIPROT:Q9YU3; UNIPARC:UPI00000F479C; EMBL:AB014368; NID:G3551304; I
A:Experimental source: Japanese patient with hepatocellular carcinoma isolate 09D09HCC
C:Genetics:
A:Gene: P
A:Introns: 303/3
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase
Alignment Scores:
Pred. No.: 3.7e-190 Length: 827
Score: 2678.50 Matches: 504
Percent Similarity: 95.91% Conservative: 12
Best Local Similarity: 93.68% Mismatches: 18
Query Match: 45.34% Indels: 5
DB: 2 Gaps: 1
US-10-761-006a-1 (1-3215) x T13468 (1-827)
QY 19 AGCTCTGCTAGATCCAGGGTGAGGGG-----CCTATATTTTCTGCTGGTGC 69
Db 291 ThrSerLysArgGlnSerSerSerGlyHisAlaValGluLeuHisProCysTrpTrpLeu 310
QY 70 CAGTTCGCGAACAGTAAACCCCTGTTCCGACTACTGCTCTCCCATATCGTCAATCTTCTC 129
Db 311 GlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsnLeuLeu 330
QY 130 GAGGACTGGGACCTTCGACCCAGAACATGGAGAACACATCAGGATTCCTAGGACCCCT 189
Db 331 GluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArgThrPro 350
QY 190 GCTGCTGTACAGGGGGGGTCTTCTCGTTGACAGAGTCTCACAATAACCCGACAGTCT 249
Db 351 AlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThrGluSer 370
QY 250 AGACTCTG-GTGCATCTTCTCTCAATTTTCTAGGGGAGCACCCACGCTGTTCCTGGCCAAA 308
Db 371 ArgLeuValValAspPheSerGlnPheSerArgLysThr**Val-SerTrpProly 390
QY 309 ATTGCGAGTCCCAACCTTCCAAATCATCATCAACCTCTTGTCTCCTCAATTTGCTGGCT 368
Db 390 sPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSerTrple 410
QY 369 ATCGCTGGATGTCTGGCGGCTTTTATCATATTCCTTCTTCTGCTGCTATGCTCTCA 428
Db 410 uSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMetProHi 430
QY 429 TCTTCTTGTGTGTTCTTCTGGGACTACCAAGGTATGTGCGCGTCTTCTCTCTACTTCCAG 488

```
Db 430 sLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerThrSerAr 450
Qy 489 GAACATCAACACACAGCAGCGGCCATGCAAGACCTGCAGACGACTCCCTGCTCAAGAAACT 548
Db 450 gAsn***AsnTyrGlnHisGlyThrMetGlnAapLeuHisGlySerCysSerArgAenLe 470
Qy 549 CTACGTTTCCCTCTGTGTGTACAAAACCTTCGGACGGAACCTCGGACGAACTGCTGATTCCTCC 608
Db 470 uTyValSerLeuLeuLeuLeuTyValThrPheGlyArgLeuHisLeuTySerHi 490
Qy 609 TCCCATCATCTCGGCTTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCGCTTCTCTCT 668
Db 490 eProlleLeuGlyPheArgGlylleProMetGlyValGlyLeuSerProPheLeuLe 510
Qy 669 GGCTCAGTTTACTAGTCGCAATTTGTCAGTGGTTCGTAGGCTTTCCTCCACCTGTTTGGC 728
Db 510 uAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHisCysLeuAl 530
Qy 729 TTTTCAGTTATATGATGATGGTATTGGGGCGGAAGTCTGTACAAATCTTGTAGTCCCT 788
Db 530 aPheSerTyrMetAspAspValValLeuGlyAlaIleValGlnHisLeuGluSerLe 550
Qy 789 TTTTACCTCTATTACCAATTTCTTTTGTCTTGGGTATACATTTAAACCCCTAATAAAC 848
Db 550 uPheThrAlaIleThrAenPheLeuLeuSerLeuGlylleHisLeuAenProHisLysTh 570
Qy 849 CAACAGTTGGGCTACTCCCTTAATCTTCATGGGATATGTAATCGAAGTTGGGTACTTT 908
Db 570 rLysArgTrpGlyTyrSerLeuAenPheMetGlyTyrValIleGlySerTrpGlyThrLe 590
Qy 909 ACCCGAGCAATATGTTACTAAACCTCAACCAATGTTTTCGAAACCTGCTGTAAATAG 968
Db 590 uProlGlnHisIleValGlnLysIleLysIleGlnCysPheArgLysLeuProValAenAr 610
Qy 969 ACCTATTGATTGGAAGTATGCTCAAGAAATTTGGGTCTTTTGGGCTTTTGTGCCCTTT 1028
Db 610 gProlleAspTrpLysValCysGln**IleValGlyLeuLeuGlyPheAlaIaProPh 630
Qy 1029 TACAAATGTGGCTATCTGCTGTGATGCTTATATGATGATGATGATGATGATGATGATG 1088
Db 630 eThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAlaLysGlnAl 650
Qy 1089 TTTTCACTTCTCGCAACTACAGGCTTCTGTGTAAACAATATCTGAACCTTTACCC 1148
Db 650 aPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHisLeuTyPr 670
Qy 1149 GTTGGCCGGCAACGGTCCGCTCTCTGCAAGTGTCTGCTGACGCAACCCCTGATG 1208
Db 670 oValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProThrGlyTr 690
Qy 1209 GGGCTTGGCCATAGCCATCAGCGCATGGCTGGAACCTTTCTGGCTCTCTGCGCATCCA 1268
Db 690 pGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLeuProIleHi 710
Qy 1269 TACTCGGAACTCTAGCAGCTTGTGTCGCGAGCGGTCTGAGCAAACTTATCGG 1328
Db 710 sThrAlaGluLeuLeuAlaIaCysPheAlaArgSerArgSerGlyAlaLysLeuIleGl 730
Qy 1329 AACCGACAACTCTGTGTCTCTCTCGAAATACACCTCTCTTTCATGGCTGCTAGGCTG 1388
Db 730 yThrAspAenSerValValLeuSerArgLysTyrThrSerPheProTrpLeuLeuGlyCy 750
Qy 1389 TGCTCCCAACTGGATCTCGCGGGAAGTCTTTGTCTTGTCTACGTCGCGCTGGAATCC 1448
Db 750 sAlaAlaAenTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLeuAenPr 770
Qy 1449 CGCGACACCGCTCTCGGCGCGTGTGGGCTCTACCGTCCCTCTTCTCATCTGCGCTT 1508
Db 770 oAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHisLeuProPh 790
Qy 1509 CGGGCGACCGCGGCGCACCTCTCTTACCGGTCTCCCGGTATGTGCTTCTCATCT 1568
Db 790 eArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSerHisLe 810
```

```
Qy 1569 GCGGACCGGTGTGCACCTTGCCTTACCTTGCACGTCGCATGAGACCCG 1620
Db 810 uProAspArgValHisPheAlaSerProLeuHisValAlaIaTrpArgProPro 827
```

RESULT 7

S71785
DNA-directed DNA polymerase (BC 2.7.7.7) - hepatitis B virus (subtype ayw, isolate patient
A; Species: hepatitis B virus, HBV
C; Variety: subtype ayw, isolate patient C1005
C; Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 09-Jul-2004
C; Accession: S71785
R; Preiller-Adams, S.; Schlayer, M.J.; Peters, T.; Hettler, F.; Gerok, W.; Rasenack, J.
submitted to the EMBL Data Library, March 1993
A; Description: Identification and sequence analysis of hepatitis B virus DNA in immunolog
A; Reference number: S32202
A; Accession: S71785
A; Molecule type: DNA
A; Residues: 1-832 <PRE>
A; Cross-references: UNIPROT:Q8JN11; UNIPROT:Q9IF40; UNIPROT:Q96846; UNIPROT:Q67892; UNIP
PROT:Q9QAG0; UNIPROT:Q9QAF3; UNIPROT:Q80X01; UNIPROT:O11885; UNIPROT:O56654; UNIPROT:O566
A; Experimental source: subtype ayw, isolate patient C1005
C; Superfamily: hepatitis virus DNA-directed DNA polymerase
C; Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:
Pred. No.: 2,42e-188 Length: 832
Score: 2654.00 Matches: 492
Percent Similarity: 95.56% Conservative: 25
Best Local Similarity: 90.94% Mismatches: 23
Query Match: 44.93% Indels: 2
DB: Gaps: 0

US-10-761-006a-1 (1-3215) x S71785 (1-832)

```
Qy 1 CTCACAACTTCCACCAAGCTCTGTAGATCCCAGGTCAGGGGCTATATTTTCCTGC 60
Db 293 LeuHisAenPheProProAenSerAlaArgSerGlnGlyGluArgProValPheProCys 312
Qy 61 TGTGGCTCCAGTCCGGAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATGCTC 120
Db 313 TrpTrpLeuGlnPheArgAenSerLysProCysSerAspTyrCysLeuSerHisIleVal 332
Qy 121 AATCTTCTCAGGACTGGGACCTCGCACCGAATCATGAGAACACACATCAGGATTCCT 180
Db 333 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisHisIleArgIlePro 352
Qy 181 AGGACCCCTCTCGTGTTCAGGCGGGTTTTCTCTGTGACAAAGATCCTCACAATACC 240
Db 353 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAenProHisAenThr 372
Qy 241 GCAGACTAGACTCTG- GTGGACTTCTCAATTTCTAGGGGAGCACCCACGTGTC 299
Db 373 AlaGluSerArgLeuValAlaAspPheSerGlnPheSerArgGlyAenHisArgVal- Se 392
Qy 300 CTGCGCAAAATTCAGTCCCAACTCCAACTCACTCAACCACTCTTGTCTCCCAATTT 359
Db 392 rTrpProLysPheAlaValProAenLeuGlnSerLeuThrAsnLeuLeuSerAenLe 412
Qy 360 GTCTCGCTACTCGCTGGATGTCTGGGGTTTTTATCATATTCCTCTTCTCTGCTGC 419
Db 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaSe 432
Qy 420 TATGCTCATCTTCTGTTGTTCTTCTGACTACCAAGGTATGTCGCCGTTTGTCTC 479
Db 432 rMetProHisLeuLeuValGlySerThrGlyLeuSerArgTyrValAlaArgValSerSe 452
Qy 480 TACTTTCAGGAACATCAACACACGACGCGGGCCATGCAAGACCTGCAAGCTCTCCTGCTC 539
Db 452 rAsnSerArgIlePheAenHisGlnArgGlyThrMetGlnAenLeuHisAspTyrCysSe 472
Qy 540 AAGGAAACTCTACGTTTCCCTCTTGTGTGCTGTACAAACCTTCGAGCGGAACTGCACCT 599
```

```
Db 472 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe 492
QY 600 GTATTCCCATCCCATCATCTCGGCTTTTCGCAAGATTCTATGGGAGTGGGCTCAGTCC 659
Db 492 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512
QY 660 GTTTCCTCCTGGCTCAGTTTACTAGTGCCTATTTGTTTCAGTGGTTCGTAGGCTTCCCA 719
Db 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 532
QY 720 CTGTTTGGCTTTTCAGTTATATGATGATGATGGTATTTGGGGCGAAGTCTGTCAACATCT 779
Db 532 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 552
QY 780 TGAGTCCCTTTTACCTCTATATACCAATTTCTTTTGTCTTTGGGATATACATTTAAACCC 839
Db 552 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 572
QY 840 TAATAAACCAACGTTGGGCTACTCCCTTAACCTCATGGGATATGTAATGGAGTTG 899
Db 572 oAsnLysThrLysArgTrpGlyTyrSerLeuHisPheMetGlyTyrValIleGlySerTy 592
QY 900 GGGTACTTTACCCAGCAACATATTGTACTTAAACTCAAGCAATGTTTTCGAAAACCTGCC 959
Db 592 rGlySerLeuProGlnAspHisIleIleGlnLysIleLysGluCysPheArgLysLeuPr 612
QY 960 TGTAATAGACCTATTGATTGGAAGATGATGCTCAAGAATTGTGGGTCTTTTGGGCTTTGC 1019
Db 612 oIleAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
QY 1020 TGCCCCCTTTTACCAATGTGGCTATCCTGCTGTATGATGCCCTTATATGATGATATACAATC 1079
Db 632 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 652
QY 1080 TAAGCAGGCTTTTCACTTTCTCGCAACTTCAAGGCTTTCTGTGTAACAATATCTGAA 1139
Db 652 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 672
QY 1140 CCTTTACCCCGTTGCCGCAACGGTCCGGTCTCTGCCAAGTGTGTCGACCAACCC 1199
Db 672 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
QY 1200 CACTGATGGGCTTGGCCATAGCCATCAGCCATGGCTGGAACTTTCTGCTCCTCT 1259
Db 692 oThrGlyTyrGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaProLe 712
QY 1260 GCCGATCCCATCTGCGAACTCCTAGCAGCTTGTGTTGTCGCGCGGCTCTCGAGCAAA 1319
Db 712 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732
QY 1320 ACTTATCGGAACCGCAACTCTGTTGCTCTCTCGGAATATACCTCTCTTCCATGGCT 1379
Db 732 nIleLeuGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProThrLe 752
QY 1380 GCTAGGCTGCTGCCAACGTGGATCTCGCGGAGCTCCTTTGTCTAGTCCCGTCGCG 1439
Db 752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 772
QY 1440 GCTGAATCCCGCGAGCAGCCGCTCGGGCCGTTTGGGGCTCTACCGTCCCTCTTCA 1499
Db 772 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyIlePheArgProLeuLeuArg 792
QY 1500 TCTGCCGTTCCGCGCAACCGGGCGCACTCTCTTTTACCGGCTTCCCGGATGTGCC 1559
Db 792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr 812
QY 1560 TTCTCATCTCCCGGACCGGTGTGCACTTCGCTTCACTCTGCACTGCGATGGAGACACC 1619
Db 812 oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832
QY 1620 G 1620
Db 832 o 832
```

RESULT 8

```
JDLVLVB
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype ayw, strain pHB320)
C:Species: hepatitis B virus, HBV
A:Note: host Homo sapiens (man)
C:Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C:Accession: A00703
R:Bichko, V.; Pushko, P.; Dreilina, D.; Pumpen, P.; Gren, E.
FEBS Lett. 185, 208-212, 1985
A:Title: Subtype ayw variant of hepatitis B virus: DNA primary structure analysis.
A:Reference number: A05237; MUID:85204397; PMID:3996597
A:Accession: A00703
A:Molecule type: DNA
A:Residues: 1-832 <BI>
A:Cross-references: UNIPROT:P03156; UNIPARC:UPI0000171088; GB:X02496; NID:962280; PIDN:CU
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase
```

```
Alignment Scores:
Pred. No.: 3,4e-188 Length: 832
Score: 2652.00 Matches: 494
Percent Similarity: 95.38% Conservative: 22
Best Local Similarity: 91.31% Mismatches: 24
Query Match: 44.90% Indels: 2
DB: 1 Gaps: 0
```

US-10-761-006A-1 (1-3215) x JDLVLVB (1-832)

```
QY 1 CTCACAACTTCCACCAAGCTCTGTAGATCCAGGGTGAGGGCGCTATATTTCTCTGC 60
Db 293 LeuHisAsnLeuProProAsnSerAlaArgSerGlnSerGluArgProValPheProCys 312
QY 61 TGGTGGCTCCAGTTCGGAAACAGTAAACCTCTTCCGACTACTGCTCTCTCCATATCGTC 120
Db 313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 332
QY 121 AATCTCTCAGAGCTGGGACCCCTGCACCAACATGGAGAACACACACATCAGGATTCCT 180*
Db 333 AsnLeuLeuGluAspTrpGlyProCysAlaGluHisGlyGluHisIleArgIlePro 352
QY 181 AGACCCCTCGCTGTTTACAGCGGGGTTTTCTCGTTGACAAAGAAATCCTCACATACC 240
Db 353 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 372
QY 241 GCAGAGCTTAGACTCTG-GTGACTTCTCTCAATTTCTAGGGGAGAGCACCCACGTGTTCC 299
Db 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnTyrArgVal-Se 392
QY 300 CTGGCCAAATTCGCAGTCCCAACCTCCAATCAGTCCCAACCTCTCTCTCCCAATTT 359
Db 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerAsnLe 412
QY 360 GTCTGGCTATCGCTGGAGTGTCTGCGCGGTTTTATCATATTTCTCTCTCATCTCTCTGC 419
Db 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 432
QY 420 TATGCTCATCTCTTGTGTTGTTCTTCTGAGTACCAAGTATGTTGCCGTTGTCCTC 479
Db 432 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 452
QY 480 TACTTCCAGAACATCAACACACGACGCGGCCATGCAAGACCTCGACGACTCCTGCTC 539
Db 452 rAsnSerArgIlePheAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 472
QY 540 AAGGAAACTCTACGTTTCCCTCTTGTGTGTGTACAAAACCTTCGAGCGAAACCTGCACCT 599
Db 472 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe 492
QY 600 GTATTCCCATCCCATCATCTCTGGGCTTTCGCAAGATTCTATGGGAGTGGGCTCAGTCC 659
Db 492 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512
```



```
QY 720 CTGTTGGCTTTCAGTTATATGATGATGTTGGGGCGAAGTCTGTACAAACATCT 779
D 532 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 552
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTGGGTATACATTTAAACCC 839
D 552 uGluSerLeuPheThrAlaValThrAsnPheLeuSerLeuGlyLysHisLeuAsnPr 572
QY 840 TAATAAAACCAACGTTGGGGCTACTCCCTTAACTTCATGGGATATGTAATGGAAAGTTG 899
D 572 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlyCysty 592
QY 900 GGGTACTTTACCCGAGCAATATTGCTATAAATCAAGCAATGTTTCGAAAACCTGCC 959
D 592 rGlySerLeuProGlnAspHisIleLeuGlnLysIleLysGluCysPheArgLysLeuPr 612
QY 960 TGTAAATAGACCTATTGATTGGAAAGTATGCTCAAGAAATTTGTGGGCTTTTGGCTTTGC 1019
D 612 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
QY 1020 TGGCCCTTTTACACAATGTGGCTATCTGCTTTGATGCTCTTTATATGTCATGTATACAATC 1079
D 632 aAlaProPheThrGlnCysGlyTyrProAlaLeuLysProLeuTyrAlaCysIleGlnSe 652
QY 1080 TAAGCAGGCTTTCACCTTTCTCGCAACTTACAAGGCTTTCTGTGTAAACAATATCTGAA 1139
D 652 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 672
QY 1140 CTTTACCCCGTTGCCGCAACGGTCCGCTCTCCCAAGTGTTCGACGCAACCC 1199
D 672 nLeuTyrProValAlaArgGlnArgProGlyLeuGlnValPheAlaAspAlaThrPr 692
QY 1200 CACTGGATGGGCTTGGCCATAGCCCATCAGCCATCGGCTGGAACTTTCTGGCTCTCT 1259
D 692 oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaProLe 712
QY 1260 GCGATCCATCTACGGAATCTCTAGCAGCTTGTGTTGTCGCGAGCGGTCTGGACAA 1319
D 712 uProIleHisThrAlaGluLeuLeuAlaLaCysPheAlaArgSerArgSerGlyAlaAs 732
QY 1320 ACTTATCGGAACCGAACACTCTGTGTCCTCTCGGAATACACCTCTCTTCCATGGCT 1379
D 732 nIleLeuGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTiple 752
QY 1380 GCTAGGCTGTGTCGCCAATCGATCTCGCGCGGACGTCTCTTTGTACGTCCTCCGCGGC 1439
D 752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTy-ValProSerAl 772
QY 1440 GCTGAATCCCGCGGACGACCCGCTCGGGCCGTTTGGGGCTCTACGTCCTCTTCTTCA 1499
D 772 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuSerArgProLeuLeuAr 792
QY 1500 TCTGCGCTTCCGCGCACCGGCGGCGACCTCTCTTTAGCGGTCTCCCGGATGTGCC 1559
D 792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr 812
QY 1560 TTTCTATCTCCGCGACCGGTGTGCACTTTCGCTTCCACCTCTGTCAGCTGCGATGGAGACCAC 1619
D 812 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832
QY 1620 G 1620
D 832 o 832
```

RESULT 10

JDVLVA

DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype ayw)

C:Species: hepatitis B virus, HBV

A:Variety: subtype ayw

C:Date: 18-Dec-1981 #sequence revision 18-Dec-1981 #text_change 09-Jul-2004

C:Accession: A00702; S20748; S53131

P:Galibert, F.; Mandart, E.; Pitoussi, F.; Tiollais, P.; Charnay, P.

Nature 281, 646-650, 1979
A:Title: Nucleotide sequence of the hepatitis B virus genome (subtype ayw) in E. coli.
A:Reference number: A93214; MUID:81012091; PMID:399327
A:Accession: A00702
A:Molecule type: DNA
A:Residues: 1-832 <GAL>
A:Cross-references: UNIPROT:P03156; UNIPARC:UPI0000129801
A:Experimental source: cloned in Escherichia coli
R:Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.; Porru, A.
A:Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negative
A:Reference number: S20745
A:Accession: S20748
A:Molecule type: DNA
A:Residues: 1-276, 'S', 278-292, 'L', 294-458, 'H', 460-465, 'N', 467-469, 'S', 471-612, 'V', 614-67;
A:Cross-references: UNIPARC:UPI00000F27F6; EMBL:X65257; NID:G59429; PIDN:CAA46352.1; PID:
A:Experimental source: subtype ayw, patient C
R:Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
A:Submitted to the EMBL Data Library, March 1995
A:Reference number: S53112
A:Accession: S53131
A:Molecule type: DNA
A:Residues: 1-16, 'D', 18-117, 'N', 119-223, 'C', 225-245, 'M', 247-271, 'Q', 273-277, 'D', 279-291 *
A:Cross-references: UNIPARC:UPI00001726S3; EMBL:X85254
C:Genetics:

A:Gene: P
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:
Pred. No.: 1,58e-187 Length: 832
Score: 2843.00 Matches: 493
Percent Similarity: 95.00% Conservative: 20
Best Local Similarity: 91.30% Mismatches: 26
Query Match: 44.74% Indels: 2
DB: 1 Gaps: 0

US-10-761-006a-1 (1-3215) x JDVLVA (1-832)

```
QY 4 CACAACTTCCACCAAGCTCTCTAGATCCAGGTGAGGGGCCTATATTTCTGCTCTGG 63
D 294 HisAsnLeuProProAsnSerAlaArgSerGlnSerGluArgProValPheProCysTrp 313
QY 64 TGGCTCCAGTTCGGNACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTCAAT 123
D 314 TrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerLeuIleValAsn 333
QY 124 CTTCTCGAGACTGGGGACCCCTGCACCGAACAATGGAGAACACAAATCAGGATTCCTAG 183
D 334 LeuLeuGluAspTrpGlyProCysAlaGluHisGlyGluHisIleArgIleProArg 353
QY 184 ACCCTGCTCGTTTACAGCGCGGGTTTTCTCGTTGACAAAGATCCTCACAAATCCGCA 243
D 354 ThrProSerArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrAla 373
QY 244 GAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTTAGGGGAGACCCACGTTCTCTG 302
D 374 GluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnTyrArgVal-SerTr 393
QY 303 GCCAAATTCGAGTCCCAACCTCCAATCACTCAACCAACCTCTTGTCTCTCAAAATTTGTC 362
D 393 pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe 413
QY 363 CTGGCTATCGCTGATGTCTCTGCGCGGTTTTATCATATTCCTTTCATCTGCTGCTAT 422
D 413 rTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAlaMe 433
QY 423 GCCTCATCTCTTGTGTTGTTCTCTTCTGGAATACCAAGGTATGTTGCCGTTTCTCTAC 482
D 433 tProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSerAs 453
QY 483 TTCCAGGAACATCAACCAACGACGGGGCCATGCAAGACCTGCACGACTCTCTGCTCAAG 542
```



```
Db 453 nSerArgIleLeuAasnAasnGlnHisGlyThrMetProAspLeuHisAspTyrCysSerAr 473
Qy 543 GAAACTCTAGCTTCCCTCTGTTGCTGCTACAAACCTTCGGACGGAACCTGCACTGTGA 602
Db 473 gAasnLeuTyrValSerLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLeuTyr 493
Qy 603 TTCCCATCCCATCATCTCGGCTTTCGCAAGATTCTCTATGGAGTGGGCTTCAGTCCGTT 662
Db 493 rSerHisProIleleuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh 513
Qy 663 TCTCCTGGCTCAGTTTACTAGTGCATTGTTCAAGTTCGTAGTGGCTTTCGCCCACTG 722
Db 513 eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCy 533
Qy 723 TTTGGCTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 782
Db 533 sLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGl 553
Qy 783 GTCCCTTTTACCTCTATTACCAATTTCTTTCTTTGGGTATACATTTAAACCCCTAA 842
Db 553 uSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAs 573
Qy 843 TAAACCAAAAGCTGGGCTACTCCCTTAACCTTCATGGATATGTAATGGAAGTTGGG 902
Db 573 nLysThrLysArgTrpGlyTyrSerLeuAasnPheMetGlyTyrValIleGlyCysTyrGl 593
Qy 903 TACTTTACCGGAGGACATATGTTACTAAACTCAAGCAATGTTTTCGAAAACCTGCTGT 962
Db 593 ySerLeuProGlnGluHisIleIleGlnLysIleGlyLysGlyCysPheArgLysLeuProIl 613
Qy 963 AAATAGACTTATGATGGAAGTATGTCAAAGAAATGTTGGCTCTTTGGGCTTTGTGTC 1022
Db 613 eAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl 633
Qy 1023 CCTTTTACAAATGTGCTATCTGCTTGCCTTGCCTTTATATGATGATGATGATGATGATG 1082
Db 633 aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLy 653
Qy 1083 GCAGCTTTTCACTTCTCGCAACTTACAGGCTTTCTGCTGTAACATATCTGAACCT 1142
Db 653 sGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAasnLe 673
Qy 1143 TTACCCCGTTCGCGCAACGCTCGGCTCTGCAAGTGTTCGTCAGCAACCCAC 1202
Db 673 uTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrProth 693
Qy 1203 TGGATGGGCTTGGCCATAGCCATAGCCATAGCCATAGCCATAGCCATAGCCATAGCCAT 1262
Db 693 rGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheSerAlaProLeuPr 713
Qy 1263 GATCCATCTCGGGAACCTCTAGCAGCTTGTGTTGCTCGCAGCGGCTCGGAGCAAACT 1322
Db 713 oIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaAasnIl 733
Qy 1323 TATCGAACCCGACACTCTGTGCTCTCTCGGAATACACCTCTTTTCCATGCTGCT 1382
Db 733 eIleGlyThrAspAasnSerValValLeuSerArgLysTyrThrSerPheProTrpLeuLe 753
Qy 1383 AGGGTGTGCTGCCAATCTGATCTCGCGGGGACGTCCTTTGTCTACGTCGCGTGGCGCT 1442
Db 753 uGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLe 773
Qy 1443 GAATCCCGGAGACGCTCTCGGGGCGGTTGGGCTCTACCGTCCCTCTTCTCATCT 1502
Db 773 uAasnProAlaAspAspProSerArgGlyArgLeuGlyLeuSerArgProLeuLeuArgLe 793
Qy 1503 GCCGTTCGGCGGACCAAGGGGCGACCTCTCTTTACGGGCTCTCCCGTATGTCCTTC 1562
Db 793 uProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValProSe 813
Qy 1563 TCATCTGCGGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
Db 813 rHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTTPArgProPro 832
```

RESULT 11

```
JDVLJ3
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adw, strain Indonef
C;Species: hepatitis B virus, HBV
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: F28925
R;Okamoto, H.; Tada, F.; Sakugawa, H.; Sastrosowignjo, R.I.; Imai, M.; Miyakawa, Y.; Ma
J. Gen. Virol. 69, 2575-2583, 1988
A;Title: Typing hepatitis B virus by homology in nucleotide sequence: comparison of surfa
A;Reference number: JS0253; MUID:89010694; PMID:3171552
A;Accession: F28925
A;Molecule type: DNA
A;Residues: 1-843 <OKA>
A;Cross-references: UNIPROT:P17393; UNIPARC:UPI00001297F8; GB:D00331; NID:g221499
C;Superfamily: hepatitis B virus DNA-directed DNA polymerase
C;Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:
Pred. No.: 1.58e-187 Length: 843
Score: 2643.00 Matches: 491
Percent Similarity: 95.38% Conservative: 25
Best Local Similarity: 90.76% Mismatches: 24
Query Match: 44.74% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1 (1-3215) x JDVLJ3 (1-843)
Qy 1 CTCACAACTTCACCAACCTCTGTAGATCCAGGTCAGGGGCTATATTTTCCTGC 60
Db 304 LeuHisValProAsnSerArgSerGlnGlySerValLeuSerCys 323
Qy 61 TGTGGCTCCAGTTCCGGAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATGCT 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerGluHisCysLeuPheHisVal 343
Qy 121 AATCTTCTCGAGGACTGGGACCTCGACCGAACATGAGAACACATCAGGATTCCT 180
Db 344 AsnLeuIleAspTrpGlyProCysAlaGluHisGlyGluHisArgThrPro 363
Qy 181 AGGACCTCTGCTGTATACAGCGGGGTTTTCTGTTGACAGAATCTCACAATACC 240
Db 364 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 383
Qy 241 GCAGACTAGACTCTG-GTGGACTCTCAATTTCTAGGGGAGCAGCCACGTGTC 299
Db 384 SerGluSerArgLeuValAlaAspPheSerGlnPheSerArgGlyAsnThrArgVal-Se 403
Qy 300 CTGGCCAAAATTCGACGTCCCAACCTCCCAATCACTCACCACCTCTGTCTCCAAATT 359
Db 403 rTTPProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAspLe 423
Qy 360 GTCCTGCTATCGCTGATGTCTGGCGGTTTTATCATATTTCTTCTTCTCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 443
Qy 420 TATGCTCATCTCTCTGTTGTTCTCTGGACTACCAAGGTATGTGGCCGTTGTCTCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 463
Qy 480 TACTTCCAGGAACATCAACCAACGACGCGGGCCATGCAAGACCTGCAGACTCTGCTC 539
Db 463 rAsnSerArgIleIleAsnHisGlnHisArgThrMetGlnAsnLeuHisAspSerCysSe 483
Qy 540 AAGAAACTCTACGTTTCCCTCTGTTGCTGTGTATACAAAACCTTCGAGCGGAACCTG 599
Db 483 rArgAsnLeuTyrValSerLeuMetLeuLeuTyrLysThrTyrGlyArgLysLeuHisLe 503
Qy 600 GTATTTCCCATCCCATCTCTGCGGCTTTCGCAAGATTCCTATGGAGTGGCCTCAGTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy 660 GTTTCTCTGCTCAGTTTACTAGTGCATTTGTTTCTAGTGGTTCGTAGGGCTTTCCCCCA 719
```

523	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi	543
720	CTGTTTGGCTTTACGTTATATGGATGATGTGGTATTGGGGCGGAAGTCTGTACAAACATCT	779
543	sCysLeuAlaPheSerTyMetAspAspValValLeuGlyAlaLysSerValGlnHisLe	563
780	TGAGTCCTCTTTTACCTCTATTACCAATTTCTTTTCTCTTTGGGTATACATTTAAACCC	839
563	uGluSerLeuTyAlaAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr	583
840	TAATAAAACAAACGTTGGGGCTACTCCCTTAACCTTCATCGGATATGTAATTCGAAGTTG	899
583	oGlnLysThrLysArgTrpGlyTySerLeuAsnPheMetGlyTyValIleGlySerTr	603
900	GGGTACTTTTACCGCAGGAACATATGTACTAAAACTCAAGCAATGTTTTCGAAAACCTGCC	959
603	pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr	623
960	TGTAAATAGACCTTATTGGAAGATGTCAAAGAAATTGTGGGTCTTTTGGCTTTGGC	1019
623	oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl	643
1020	TGCCCTTTTACACAATGTGGCTATCTCGCTTGATGCCTTTATATGCATGTATACAATC	1079
643	aAlaProPheThrGlnCysGlyTyTrpAlaLeuLeuMetProLeuTyAlaCysIleGlnAl	663
1080	TAGCAGGCTTTCACTTTCTCGCAACTTACAGGCCCTTTCTGTGTTAAACAATATCTGAA	1139
663	aLysGlnAlaPheThrPheSerProThrTyTrpLysAlaPheLeuAsnLysGlnTyLeuAs	683
1140	CTTTTACCCTGTGSCCGGCAACGGTCCGGTCTCTGCAAGTGTTCCTCAGCGAACCC	1199
683	nLeuTyPProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr	703
1200	CACGTGATGGGCTTGCCATAGCCCATCAGCGCATGGCTGGAACTTTCTGGCTCTCTCT	1259
703	oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValSerProLe	723
1260	GCGCATCTACTCGGAACCTCTAGCAGCTGTGTTGCTCGAGCGGCTCGAGCAAA	1319
723	uProIleHisThrValGluLeuLeuAlaIaCysPheAlaArgSerArgSerGlyAlaLys	743
1320	ACTTATCGGAACCGCAACCTCTGTGTCTCTCTCGGAATACACCTCTCTTCATGGCT	1379
743	sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLe	763
1380	GCTAGGTGTGTCGCAACTGGATCTCGCGGGAGCTCTTTGTCTACGTCGCGTCGGC	1439
763	uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAl	783
1440	GCTGAATCCCGGACGACCCGCTCTCGGGCCGTTTGGGCTCTACCGTCCCTCTTCTCA	1499
783	aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyArgProLeuLeuAr	803
1500	TCTGCCGTTCCGCGCACCGGCGCACTCTCTTTTACCGGCTCTCCCGCTATGTGCC	1559
803	gLeuProTyArgProThrThrGlyArgThrSerLeuTyAlaAspSerProSerValPr	823
1560	TTCTCATCTCCGGACCGTGTGACCTTCGCTTCACCTCTCAGCTGCATGGAGACCACC	1619
823	oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaIaTrpArgProPr	843
1620	G 1620	
843	O 843	

RESULT 12

JDVLA1
KESOL

C:Species: hepatitis B virus, HBV
C:Species: hepatitis B virus (strain alpha)

C/species: hepatitis B virus, HBV
C/Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text change 09-Jul-2004

C:Accession: C34773

R;Tong, S.; Li, J.; Vitvitski, L.; Trepo, C.
Virology 176, 596-603, 1990
A;Title: Active hepatitis B virus replication in the presence of anti-HBe is associated with
A;Reference number: A34773; MUID:90266476; PMID:2345966
A;Accession: C34773
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-832 <TON>
A;Cross-references: UNIPROT:P24024; UNIPARC:UPI00001297F7; EMBL:M32138; NID:G329667; PID:
C;Genetics:
A;Gene: P
C;Superfamily: hepatitis virus DNA-directed DNA polymerase
C;Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:			
Pred. No.:	1.87e-187	Length:	832
Score:	2642.00	Matches:	492
Percent Similarity:	94.82%	Conservative:	21
Best Local Similarity:	90.94%	Mismatches:	27
Query Match:	44.73%	Indels:	2
DB:	1	Gaps:	0

UUS-10-761-006A-1 (1-3215) x JPDVLA1 (1-832)

Qy	1	CTCCACAACATCTCCACAAGCTCTGCTAGATGCCAGGGTGAGGGCCCTATATTTTCTCTGC	60
Db	293	LeuHisAsnLeuProProAsnSerAlaArgSerGlnSerGluArgProValSerProCys	312
Qy	61	TGTTGGCTCCAGTTCGGGAACAGTAACCCCTGTTCCGACTACTGCCTCTCCCATATCGTC	120
Db	313	TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal	332
Qy	121	AATCTTCTCGAGGACTGGGGACCTCGACCCGAAATGAGAGAACACAAATCAGGATTCTCT	180
Db	333	AsnLeuLeuGluAspTrpGlyProCysAlaGluHisGlyGluHisIleArgIlePro	352
Qy	181	AGGACCCCTGCTCGTGTGTACAGGGGGGGTTTTTCTCGTTGACAAGAAATCTCCACAATACC	240
Db	353	ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr	372
Qy	241	GCAGACTAGACTCTG- GTGGACTTCTCTCAATTTTCTAGGGGGAGACCCACGCTGTC	299
Db	373	GluGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnHisArgVal- Se	392
Qy	300	CTGGCCAAAATTCGCAGTCCCCCACTCCAACTCACTACCAACACTTGTGTCTCCCAATTT	359
Db	392	rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnIle	412
Qy	360	GTCCTGCTATCGCTGGATGTGCTGGGGCGTTTTTATCATATTCCTTTCATCTCTGCTGC	419
Db	412	uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl	432
Qy	420	TATGCCTCATCTCTGTTGGTTCTTCTGGACTACCAAGGTATGTTGCCGTTTGTCTCTC	479
Db	432	aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe	452
Qy	480	TACTTCCAGGAACATCAACCACGACGGGGCCCATCGAACCTGCACGACTCGCTCGTC	539
Db	452	rAspSerArgIlePheAsnHisGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe	472
Qy	540	AAGGAAACTCTACGTTTTCCCTTGTGTGCTGTACAAAACCTTCGGACGGAAACTGCATT	599
Db	472	rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrGlnThrPheGlyArgGlyLeuHisIle	492
Qy	600	GTATTCCTCCATCCATCTCGGCTTTTCGAAGATTCTTATGGAGTGGGCTCTCAGTCC	659
Db	492	uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr	512
Qy	660	GTTTCTCTGGCTCAGTTTACTAGTGCATTGTGTTCAGTGGTTCGTAGGGCTTTCCCCCA	719
Db	512	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi	532
Qy	720	CTGTTTGGCTTTTCAGTTTATATGGATGATGTGGTATTTGGGGCGCGAAGTCTGTACAACTCT	779

```
Db 532 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysThrValHisLe 552
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGCTCTTGGGTATATTAACCC 839
Db 552 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 572
QY 840 TAATAAACCAAAAGTTGGGGCTACTCCCTTAACCTTCATGGGATATGTAATTCGAAGTTG 899
Db 572 oAenLysThrLysArgTrpGlyTyrSerLeuHisPheMetGlyTyrValIleGlyCystY 592
QY 900 GGGTACTTTACCGCAGGAACATATGCTACTAAACCTCAAGCAATGTTTTCGAAAATGCC 959
Db 592 rGlySerLeuProGlnAspHisIleIleGlnLysIleLysGluCysPheArgLysLeuPr 612
QY 960 TGTAAATAGACCTATTGATTTGGAAGTATGTCGGAAGTATGTCGCTTTTGGCTTTGC 1019
Db 612 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
QY 1020 TGCCCTTTTACAAAGTGGCTATCTGCTTGTATGCTTATATGATGCTATATCAATC 1079
Db 632 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 652
QY 1080 TAAGCAGCTTTTCACTTTCTCGCCAACTTACAAGCCCTTTCTGTGTAAACAATATCTGAA 1139
Db 652 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 672
QY 1140 CTTTATCCCTTGGCCGCAACGGTCCGCTCTCTGCAAGTGTCTGCTGACGCAACCCC 1199
Db 672 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
QY 1200 CACTGGATGGGCTTGGCCATAGCCATCAGCCATGCTGGAACCTTTCTGCTCTCTCT 1259
Db 692 oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheGlnAlaProLe 712
QY 1260 GCCATCCATACTCGGAACCTCTAGCAGCTTGTGCTCGCAGCGGCTCGGAGCAAA 1319
Db 712 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732
QY 1320 ACTTATCGGAACCGAACCTCTGTTGCTCTCTCGGAATACACCTCTCTTCCATGGCT 1379
Db 732 nIleLeuGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 752
QY 1380 GCTAGGTGTCTCCAACTGGATCTCGCGGAGCTCTTGTCTACGTCCTCGCTCGGC 1439
Db 752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 772
QY 1440 GCTGAATCCGCGCAGCAGCCGCTCTCGGGCGGTTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db 772 aLeuAsnProAlaAspAspProSerArggLyArgLeuGlyLeuSerArgProLeuLeuAr 792
QY 1500 TCTCGCTTCCGCGCCGACCGCGGCGCACCTCTCTTTTACGCGTCTCCCGGTATGTC 1559
Db 792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr 812
QY 1560 TTCTCATCTCGCGACCGTGTGCACTTGTCTTCCCTCTGACCTCGCATCGCATGAGACCA 1619
Db 812 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832
QY 1620 G 1620
Db 832 o 832
```

RESULT 13

S20757
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype ayw, patient B)
C:Species: hepatitis B virus, HBV
A:Variety: subtype ayw, patient E
C:Date: 20-Feb-1995 #sequence_revision 22-Nov-1996 #text_change 09-Jul-2004
C:Accession: S20757
R:Iai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.; Porru, A.
submitted to the EMBL Data Library, March 1992
A:Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negati

A:Reference number: S20745
A:Accession: S20757
A:Molecule type: DNA
A:Residues: 1-832 <LAI>
A:Cross-references: UNIPROT:Q67882; UNIPARC:UPI00000F25B1; EMBL:X65259; NID:G59439; PIDN:
A:Experimental source: subtype ayw, patient E
C:Genetics:
A:Gene: P
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores: 4.38e-187 Length: 832
Pred. No.: 2637.00 Matches: 494
Score: 95.19% Conservative: 21
Percent Similarity: 91.31% Mismatches: 25
Best Local Similarity: 44.64% Indels: 2
Query Match: 1
DB: 0

US-10-761-006A-1 (1-3215) x S20757 (1-832)

```
QY 1 CTCCACAACATTCACCAAGCTCTGCTAGATCCAGGCTGAGGGGCTATATTTCTCTGC 60
Db 293 LeuHisAsnLeuProAsnSerAlaArgSerGlnSerGluArgProValPheProCys 312
QY 61 TGGTGGCTCCAGTTCGCGAACAGTAAACCTGTTCCGACTTACTGCTCTCCCATATGTC 120
Db 313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerLeuIleVal 332
QY 121 AATCTTCTCGAGACTGGGGACCTCCACCGAACATCGAGAACACAAACATCATGATTCCT 180
Db 333 AsnLeuArgGluAspTrpGlyProCysThrGluHisGlyGluHisIleArgIlePro 352
QY 181 AGACCCCTGCTCGTGTACAGCGGGGTTTCTCGTTCGACAGAAATCTCAACAATACC 240
Db 353 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 372
QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGACCCACGTTCTC 299
Db 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnTyrArgVal-Se 392
QY 300 CTGGCCAAATTCGAGTCCCAACCTCCCAATCATCATCACCACCTCTTGTCTCCCAATTT 359
Db 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 412
QY 360 GTCTGGCTATCTCGATGTGTCTCGGGGTTTATCATATTCCTCTCTCATCTCTGCTGC 419
Db 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 432
QY 420 TATGCTCATCTTCTTGTGTTCTTCTGAGCTTACCAAGGTATGTTGCCGTTTGTCTCTC 479
Db 432 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 452
QY 480 TACTTCCAGGAACATCAACCAACGACGCGGGCCATGCAAGACCTGCACGACTCTCTGCTC 539
Db 452 rAsnSerArgIlePheAsnAsnGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 472
QY 540 AAGAAACTCTAGTTTCCCTCTTGTGTCTGTACAAAACCTTCGGACGGAACTGCACATT 599
Db 472 rArgAsnLeuTyrValSerLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe 492
QY 600 GTATTTCCCATCATCATCTCTGCGCTTTCGCAAGATTCTCTATGGAGTGGGCTCAGTCC 659
Db 492 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512
QY 660 GTTTCCTCTGGCTCAGTTTACTAGTGCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCCA 719
Db 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 532
QY 720 CTGTTTGGCTTTTCAGTTATATATGATGTGTTTTCGGGGCGGAGTCTGTACACATCT 779
Db 532 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 552
```



```
|||||.....|
552 uGluSerIleTyrThrAlaValThrAnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 572
Qy 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACCTCATGGGATATATAATGGAAGTTG 899
Db 572 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrIleGlySerTr 592
Qy 900 GGGTACTTTACCGCAGGACATATGTTACTAAACCTCAAGCAATGTTTTCGAAAACGTGCC 959
Db 592 pGlySerLeuProGlnAspHisIleValGlnLysIleGlnCysPheArgLysLeuPr 612
Qy 960 TGTAATAGACCTATTGATTGGAAGATGTCACAAAGAAATTGTGGGTCTTTTGGGCTTTGC 1019
Db 612 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
Qy 1020 TGCCCCCTTTTACAAATGTGGCTATCTGCTTATGCTTATGATGCTTATGATGATGATGATG 1079
Db 632 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAl 652
Qy 1080 TAAGCAGGCTTTCATCTTCTCGCAACTTACAAAGGCTTTCGTGTAACAATATCTGAA 1139
Db 652 aLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuArgThrGlnTyrLeuTh 672
Qy 1140 CCTTTACCCCTTCCCGGCAACGCTCCGGTCTCTGCCAAGTGTTCGTGACGCAACCC 1199
Db 672 rLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
Qy 1200 CACTGGATGGGCTTGGCCATAGCCCATACGCGCATGGCTGGAACTTTCGGTCTCTCT 1259
Db 692 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 712
Qy 1260 GCCGATCCATATCTCGGAACCTTAGCAGCTTGTTCCTCGCAGCGGCTCGGAGCAA 1319
Db 712 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732
Qy 1320 ACTTATCGGAACCGACACTCTGTTGCTCTCTCGGAAATACACCTCTCTTTCATGGCT 1379
Db 732 nIleIleGlyThrAspAsnSerValValLeuSerProLysTyrThrSerPheProTrpLe 752
Qy 1380 GCTAGGTTGTGTCGCAACTGGATCCTCGCGGGACGTCCTTTGTCTACGTCCTCGGC 1439
Db 752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgArgThrSerPheValTyrValProSerAl 772
Qy 1440 GCTGAATCCCGGACGACCGCTCTCGGGGCGGTTGGGCTCTACCGTCCCTTCITCA 1499
Db 772 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuAr 792
Qy 1500 TCTGCCGTTCCGGCGCACCGGCGGACCTCTCTTTACGGGGTCTCCCGTATGTGCC 1559
Db 792 gProTrpPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 812
Qy 1560 TTCTCATCTGCGGACCGTGTGCTACCTTCGCTTTCACCTCTGCACTGCGATGAGACACC 1619
Db 812 oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832
Qy 1620 G 1620
Db 832 o 832
RESULT 15
JDVLJ1
DNA-directed DNA polymerase (BC 2.7.7.7) - hepatitis B virus (subtype adw, strain Japan/
C:Species: hepatitis B virus, HBV
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C:Accession: D28925
R:Okamoto, H.; Tsuda, F.; Sakugawa, H.; Saetosoewignjo, R.I.; Imai, M.; Miyakawa, Y.; M
J. Gen. Virol. 69, 2575-2583, 1988
A:Title: Typing hepatitis B virus by homology in nucleotide sequence: comparison of surd
A:Reference number: J50253; MUID:89010694; PMID:3171552
A:Accession: D28925
A:Molecule type: DNA
A:Residues: 1-843 <OKA>
A:Cross-references: UNIPROT:P17394; UNIPARC:UPI00001297F9; GB:D00329; NID:G221497
```

C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores: 2.04e-186 Length: 843
Pred. No.: 2628.00 Matches: 489
Score: 94.27% Conservative: 21
Best Local Similarity: 90.39% Mismatches: 30
Query Match: 44.49% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1 (1-3215) x JDVLJ1 (1-843)

```
Qy 1 CTCACAAACATTCACCAAGCTCTGTAGATCCAGGGTGAAGGGCCCTATATTTCTCTGC 60
Db 304 LeuHisHisPheProAsnSerSerArgSerGlnSerGlnGlySerValLeuSerCys 323
Qy 61 TGTGGCTCCAGTTCGGAAACAGTAACCTGTTCCGACTACTGCTCTCCCATATGCTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerGluTyrCysLeuSerHisIleVal 343
Qy 121 AATCTTCTCAGACTGGGACCTCTGCACCGAACATCGAGAACACACATCAGGATTCCT 180
Db 344 AsnLeuIleGluAspTrpGlyProCysAlaGluHisGlyGluHisArgIleArgThrPro 363
Qy 181 AGGACCCCTCTCTGTGTACAGCGGGGTTTTCTCTGTGACAAGAAATCTCACAATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
Qy 241 GCAGAGCTTAGACTCTG- GTGGACTTCTCTCAATTTCTAGGGGAGACCCACGTTCTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnThrArgVal-Se 403
Qy 300 CTGGCCAAATTCGACGTCCTCCCACTCCCAATCACTCACTCAACCACTCTGTCTCCAAATT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
Qy 360 GTCTGGCTACTCGTGTGATGTCTCGGGGTTTTTATCATATTTCTCTTCTCTCTCTGCTC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 443
Qy 420 TATGCTCTCATCTCTCTGTTGGTCTCTCTGACTACCAAGGTATGTTGCCCTTTGCTCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 463
Qy 480 TACTTCCAGGAACATCAACACACGACGCGGCCCATCGACACCTCGACACCTCTCTGCTC 539
Db 463 rAsnSerArgIleIleAsnHisGlnHisGlyThrMetGlnAspLeuHisAsnSerCysSe 483
Qy 540 AAGGAAACTCTACCTTTCCCTCTTGTGTGTGTATACAAAACCTTCGACGCGGAACTGCACTT 599
Db 483 rArgAsnLeuTyrValSerLeuMetLeuLeuTyrLysThrTyrGlyTrpLysLeuHisLe 503
Qy 600 GTATTCCCATCCCATCATCTCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy 660 GTTTCCTCTGCTCAGTTTACTAGTCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHI 543
Qy 720 CTGTTTGGCTTTCTAGTTATATGATGTGTTGGGGGCGGAGTCTGTACAACTCT 779
Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
Qy 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACCC 839
Db 563 uGluSerLeuTyrAlaAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
Qy 840 TAATAAAACCAAACTTGGGCTCTCTCCCTTAACCTTCATGGATATGATTAATGGAAGTTG 899
Db 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
```

```

QY 900 GGGTACTTTACCGAGGAACATATTGTACTAAACTCAAGCAATGTTTTTCGAAAACTGCC 959
Db ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
603 PGlyThrTrpProGlnAspHisIleValGlnAsnPheLeuCysPheArgLysLeuPr 623
QY 960 TGTAATAGACCTATTGATTGGAAGTATGTCAAAGAAATTGTGGGTCTTTTGGGCTTTGC 1019
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCTTTTACACAATGTGGCTATCCTGCTGTATGCTCTTTATATGCATGTATACAATC 1079
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAl 663
QY 1080 TAAGCAGGCTTTTCACTTCTCGCAACTTACAAAGGCTTTCTGTGTAACAATATCTGAA 1139
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
663 aLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuSerLysGlnTyrMetTh 683
QY 1140 CCTTTACCCCGTTGCCCGGCAACGGTCTCTGCCAAGTCTTTGCTGACGCAACCCCC 1199
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
683 rLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 703
QY 1200 CACTGGATGGGGCTTGGCCATAGCCCATAGCCGATGGCTGGAACTTTCTGGCTCTCT 1259
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
703 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValSerProLe 723
QY 1260 GCCGATCCATACATGCGGAACCTCTAGCAGCTTCTTTTGTCTCGCAGCGGTCTGGAGCAA 1319
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
723 uProIleHisThrAlaGluLeuAlaIleCysPheAlaArgSerArgSerGlyAlaAs 743
QY 1320 ACTTATCGGAACCGACAACCTCTGTCTCTCTCGGAATACACCTCTTCCATGGCT 1379
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
743 nLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
QY 1380 GCTAGGTTGTGCTGCCAACCTGGATCCTGCGCGGAGCTCTTTGTCTAGCTCCCGTCGGC 1439
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
QY 1440 GCTGAATCCCGGAGCAGCCGCTCGGGCGCTTTGGGGCTCTACCGTCCCGCTCTTCA 1499
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuAr 803
QY 1500 TCTGCGGTTCCGCGGACCAACCGGGCGCACCTCTCTTTACGCGGTCTCCCGCTATGTGC 1559
Db ||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
803 gLeuProTyrArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr 823
QY 1560 TTCTCATCTGCGGAGCCGTGTGACCTTCGCTTCACTCTGACGTGCGATGGAGACCACC 1619
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
QY 1620 G 1620
Db 843 o 843

```

Search completed: December 2, 2005, 00:02:30
Job time : 160 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 1, 2005, 23:14:50 ; Search time 706.5 Seconds
(without alignments)
6421.156 Million cell updates/sec

Title: US-10-761-006A-1
Perfect score: 5907
Sequence: 1 CTCACACATTCACCAAG.....CCTCAGGCCACGAGTGGAA 3215

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US10761006/runat.01122005.113952.26869/app.query.fasta.1.3399
-DB=UniProt -OPMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORES=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10761006.CGN.1.1199 @runat.01122005.113952.26869 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt.05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4130	69.9	843	Q69616_HPBVO	Q69616 hepatitis b
2	3745.5	63.4	832	Q67907_HPBVO	Q67907 hepatitis b
3	3693.5	62.5	842	Q69590_HPBVO	Q69590 hepatitis b
4	2824	47.8	843	Q7TD85_HPBVO	Q7TD85 hepatitis b
5	2822	47.8	843	Q4FD83_HPBVO	Q4FD83 hepatitis b
6	2810	47.6	540	Q69026_HPBVO	Q69026 hepatitis b
7	2808	47.5	843	Q69026_HPBVO	Q69026 hepatitis b
8	2807	47.5	843	Q69026_HPBVO	Q69026 hepatitis b
9	2806	47.5	540	Q69026_HPBVO	Q69026 hepatitis b
10	2806	47.5	843	Q69026_HPBVO	Q69026 hepatitis b
11	2804	47.5	843	Q69026_HPBVO	Q69026 hepatitis b
12	2803	47.5	843	Q69026_HPBVO	Q69026 hepatitis b
13	2803	47.5	843	Q69026_HPBVO	Q69026 hepatitis b
14	2802	47.4	838	Q69026_HPBVO	Q69026 hepatitis b
15	2802	47.4	843	Q69026_HPBVO	Q69026 hepatitis b
16	2802	47.4	843	Q69026_HPBVO	Q69026 hepatitis b

17	2801	47.4	843	2	Q80MR4_HPBVO	Q80MR4 hepatitis b
18	2800	47.4	843	2	Q5DVZ8_HPBVO	Q5DVZ8 hepatitis b
19	2799	47.4	843	2	Q5KE35_HPBVO	Q5KE35 hepatitis b
20	2799	47.4	843	2	Q4FD61_HPBVO	Q4FD61 hepatitis b
21	2798	47.4	843	2	Q91549_HPBVO	Q91549 hepatitis b
22	2798	47.4	843	2	Q80H07_HPBVO	Q80H07 hepatitis b
23	2798	47.4	843	2	Q5KR19_HPBVO	Q5KR19 hepatitis b
24	2797	47.4	843	2	Q6YLM0_HPBVO	Q6YLM0 hepatitis b
25	2796	47.3	843	2	Q80MQ8_HPBVO	Q80MQ8 hepatitis b
26	2796	47.3	843	2	Q8JVC9_HPBVO	Q8JVC9 hepatitis b
27	2795	47.3	837	2	Q5SDK1_HPBVO	Q5SDK1 hepatitis b
28	2794	47.3	843	2	Q9QN54_HPBVO	Q9QN54 hepatitis b
29	2794	47.3	843	2	Q9YZR5_HPBVO	Q9YZR5 hepatitis b
30	2794	47.3	843	2	Q8BCI7_HPBVO	Q8BCI7 hepatitis b
31	2792	47.3	843	2	Q91547_HPBVO	Q91547 hepatitis b
32	2792	47.3	843	2	Q99217_HPBVO	Q99217 hepatitis b
33	2792	47.3	843	2	Q9QM17_HPBVO	Q9QM17 hepatitis b
34	2792	47.3	843	2	Q9QMJ8_HPBVO	Q9QMJ8 hepatitis b
35	2791.5	47.3	844	2	Q4FD77_HPBVO	Q4FD77 hepatitis b
36	2791	47.2	540	2	Q97976_HPBVO	Q97976 hepatitis b
37	2791	47.2	843	2	Q39671_HPBVO	Q39671 hepatitis b
38	2791	47.2	843	2	Q81134_HPBVO	Q81134 hepatitis b
39	2791	47.2	843	2	Q8VIM4_HPBVO	Q8VIM4 hepatitis b
40	2790	47.2	843	2	Q91541_HPBVO	Q91541 hepatitis b
41	2790	47.2	843	2	Q5SDK5_HPBVO	Q5SDK5 hepatitis b
42	2790	47.2	843	2	Q5KR43_HPBVO	Q5KR43 hepatitis b
43	2790	47.2	843	2	Q5R219_HPBVO	Q5R219 hepatitis b
44	2790	47.2	843	2	Q9QMK4_HPBVO	Q9QMK4 hepatitis b
45	2789	47.2	843	2	Q9PW18_HPBVO	Q9PW18 hepatitis b

ALIGNMENTS

RESULT 1

Q69616_HPBVO

ID Q69616_HPBVO PRELIMINARY; PRT; 843 AA.

AC Q69616;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Polymyrase (Fragment).

GN Name=P;

OS Hepatitis B virus.

OC Viruses; Retro-transcribing viruses; Hepadnaviridae;

OC Orthohepadnavirus.

OK NCBI_taxid=10407;

FN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Genotype C;

RX MEDLINE=94120723; PubMed=8291231;

RA Norder H., Courouce A.M., Magnus L.O.;

RT "Complete genomes, phylogenetic relatedness, and structural proteins

of six strains of the hepatitis B virus, four of which represent two

new genotypes.";

RL Virology 198:489-503 (1994).

FN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Genotype c;

RA Norder H.M.L.;

RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.

DR EMBL: X75656; CAA53338.1; -; Genomic_DNA.

DR GO: GO:0003677; F:DNA binding; IEA.

DR GO: GO:0003887; F:DNA-directed DNA polymerase activity; IEA.

DR GO: GO:0004523; F:Ribonuclease H activity; IEA.

DR GO: GO:0003723; F:RNA binding; IEA.

DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.

DR GO: GO:0006278; P:RNA-dependent DNA replication; IEA.

DR InterPro: IPR000462; DNAPol_viral_C.

DR InterPro: IPR000477; DNAPol_viral_N.

DR InterPro: IPR000477; RVTrse.

DR Pfam: PF00336; DNA_pol_viral_C; 1.

DR Pfam: PF00242; DNA_pol_viral_N; 1.

DR Pfam: PF00078; RVT_1; 1;
DR ProDom: PD00814; DNapi_viral_C; 1.
DR POSITE; PS00878; RT_POL; 1.
FT CHAIN 541 588 core antigen.
FT NON_TER 1 1
FT NON_TER 843 843
SQ SEQUENCE 843 AA; 94603 MW; 934BEEC1F4E235D0 CRC64;

Alignment Scores:
Pred. No.: 1 55e-299 Length: 843
Score: 4130.00 Matches: 802
Percent Similarity: 76.26% Conservative: 17
Best Local Similarity: 74.67% Mismatches: 24
Query Match: 69.92% Indels: 232
DB: 2 Gaps: 1

US-10-761-006A-1 (1-3215) x Q69616_HPBVO (1-843)

QY 1 CTCACAAATCCCAAGCTCTGTAGATCCAGGGTGGGGGCTATATTTCTCTGC 60
DB 1 LeuHisAsnIleProThrSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 20
QY 61 TGTGTGCTCCAGTTCGGGAACAGTAAACCTGTTCGGACTACTGCTCTCCCATATCGTC 120
DB 21 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisPheVal 40
QY 121 AATCTTCTCAGGACTGGGGACCTCCACCGAACATGGAGAACACATCAGGATTCCT 180
DB 41 AsnLeuLeuGluAspTrpGlyProCysAsnGluHisGlyGluHisAsnIleArgIlePro 60
QY 181 AGGACCTCTCGTGTACAGCGGGGTTTCTCGTTGACGAGATCTCTCAATACC 240
DB 61 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 80
QY 241 CGAGAGCTAGACTCG -GTGAGCTTCTCTCAATTTCTAGGGGAGCACCCACGNGTTC 299
DB 81 ThrGluSerArgLeuGlyValAspPheSerGlnPheSerArgGlySerThrLysVal -Se 100
QY 300 CTGGCCAAATTCGAGTCCCAACCTCCAATCACTCACCAACCTCTGTCTCTCCAAATTT 359
DB 100 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 120
QY 360 GTCTCGCTATCGCTGGATGTCTCGGGGTTTATCATATTCCTCTTCATCTCTCTGC 419
DB 120 uSerTrpLeuSerLeuAspValSerAlaPheTyrHisLeuProLeuHisProAlaAl 140
QY 420 TATGCTCATCTTCTTGTGTCTCTGACTACCAAGTATGTTGCCGTTTGTCTC 479
DB 140 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 160
QY 480 TACTTCCAGGAACATCAACACGACGCGGGCCATGCAAGACCTTCACGACTCTCTGCTC 539
DB 160 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 180
QY 540 AAGGAACCTACGTTTCTCTGTGTCTGTACAAACCTTCGGACGGAACCTGCACTT 599
DB 180 rArgAsnLeuTyrValSerLeuMetLeuTyrLysThrPheGlyArgLysLeuHisLe 200
QY 600 GTATTCCTCCATCCATCATCTCTGGGCTTCGCAAGATTCTATGGGAGTGGGCTCAGTCC 659
DB 200 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 220
QY 660 GTTTCCTCTCGCTCAGTTACTAGTGCATTTGTTTCAGTGGTTCGAGGGCTTCCCCCA 719
DB 220 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 240
QY 720 CTGTTTGGCTTTCAGTATATGATCATCTGTATTTGGGGGGAAGCTGTACACATCT 779
DB 240 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 260
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
DB 260 uGluSerLeuTyrThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 280

QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACCTTCATGGGATATGTAAATGGAAGTTG 899
DB 280 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 300
QY 900 GGTACTTTTACCGCAGGAACATATTTGCTACTAAACTCAAGCAATGTTTTCGAAAACCTGCC 959
DB 300 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 320
QY 960 TGTAAATAGACCTATTGATTGAAAGTATGTCAAGAATTTGTGGGTCTTTTGGGCTTTTC 1019
DB 320 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 340
QY 1020 TGCCCTTTTACCAATGTGGCTATCTCTCTCTGTATGCTTATATGATGATATACAATC 1079
DB 340 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAl 360
QY 1080 TAAGCAGGCTTTCACCTTCTGCCCACTTACAGGCTTCTGTGTAAACAATATCTGAA 1139
DB 360 aLysGlnAlaPheThrPheSerSerThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 380
QY 1140 CTTTACCCCGTTGCCGCAACGGTCCGCTCTCTGCCAAGTGTCTGTGCTGACGCAACCCC 1199
DB 380 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 400
QY 1200 CACTGATGGGCTTGGCCATAGGCCATCAGCCGATGGCTGGAACCTTTCTGGCTCTCT 1259
DB 400 oThrGlyTrpGlyLeuAlaIleGlyAsnGlnArgMetArgGlyThrPheValAlaProLe 420
QY 1260 GCGATCCATATCGGGAACCTCTAGCAGCTTGTGTTGCTCGCAGCGGCTCTGGAGCAAA 1319
DB 420 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 440
QY 1320 ACTTATCGGAACCGCAACTCTGTTGTCTCTCTCGGAATACACCTCTCTTCCATGGCT 1379
DB 440 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 460
QY 1380 GTAGGTTGTGTCCAACTGGATCTTCGCGGGAGCTCTTGTGTCTACGTCCGCTCGC 1439
DB 460 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 480
QY 1440 GTGAATCCCGGGAGCAGCCCTCTCGGGCGGTTGGGGCTCTACCGTCCCTTTCTTCA 1499
DB 480 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyIleTyrArgProLeuLeuAr 500
QY 1500 TCTGCGTTCCTCGCGCGCACCTCTCTTTACGGGCTCTCCCGGTATCTGCTCC 1559
DB 500 gLeuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 520
QY 1560 TTCTCATCTGCGGACCGGTGTCACCTTCGCTTCACCTCTGACGCTGCATGGAGACCACC 1619
DB 520 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaLysArgProPr 540
QY 1620 GTGAACGCGACCGGCTCTTGCCCAAGGCTTATATAAGAGGACTCTTGAGACTCTCAGCA 1679
DB 540 o----- 540
QY 1680 ATGTCAAACGACCGACTTGAGGCATATCTCAAAGACTGTGTGTTAAAGACTGGGAGGAG 1739
DB 540 ----- 540
QY 1740 TTGGGGGAGGAGATTAGGTTAAAGATTATATGTAAGAGGCTGTAGGCATAAATTTGGTC 1799
DB 540 ----- 540
QY 1800 TGTTTACCAGCACCACCTGCAACTTTTCTCTCTCGCTTAATCATCTCATGTTTCATGTCCTA 1859
DB 540 ----- 540
QY 1860 CTGTTTCAAGCTTCAAGCTGTGCTTGGGTGGCTTTGGGACATGGACATTGACCCCGTATA 1919
DB 540 ----- 540


```
QY 1920 AAGNAITGGAGCATCTGCTGAGTACTCTCTTTTGGCCTCTGACTTCTTCCGCTCTA 1979
Db 540 ----- 540
QY 1980 TTGAGATCTCTCGACACCGCTCTGCTCTGTATCGGAGGCTTAGAGTCTCCGGAAC 2039
Db 540 ----- 540
QY 2040 ATTGTCGCTCACCATACAGCACTCAGGCAAGCTATTTGTGTGGGTGAGTTGATGA 2099
Db 540 ----- 540
QY 2100 ATCTGGCCACCTGGGTGGGAAGTAATTGGAAGATCCAGCATCCAGGAATTAGTAGTCA 2159
Db 540 ----- 540
QY 2160 GCTATGTCAACGTTAATATATGGGCTAAACTCAGACAAATATTGTGTTTCACATTTCT 2219
Db 540 ----- 540
QY 2220 GTCTTACTTTTGAAGAGAACTGTTCTTGAGTACTTGGTATCTTTTGGAGTGTGATTC 2279
Db 540 ----- 540
QY 2280 GCACCTCTACCGCTTACAGACCACCAAAATGCCCTATCTTATCAACACTTCCGGAACCTA 2339
Db 541 -----Met--ProLeuSerTyrGlnHisPheArgLysLeu 551
QY 2340 CTGTTGTTAGACGAGGAGGCTCCCTAGAGAAGAACTCCCTCCGCTCGGACAGCAA 2399
Db 552 LeuLeuLeuAspValGluAlaGlyProLeuGluGluLeuProArgLeuAlaAspGlu 571
QY 2400 GGTCTCAATCGCCGCTCGGAGAGATCTCAATCTCGGATCTCAACGTTAGTATTCCT 2459
Db 572 GlyLeuAsnHisArgValAlaGluAspLeuAsnLeuGlyAsnProAsnValSerIlePro 591
QY 2460 TGGACTATAAGTGGGAACCTTACTGGGCTTTATCTCTACTGACCTGCTTTAAT 2519
Db 592 TrpThrHisLysValGlyAsnPheThrGlyLeuTyrSerSerThrValProValPheAsn 611
QY 2520 CCCGAGTGGCAAAATCTCTCTCTCCATTTACAGAGGACATTTATTAATPAGA 2579
Db 612 ProGlnTrpGlnThrProSerPheProAspIleHisLeuGlnGluAspIleAsnArg 631
QY 2580 TGTCACAAATATGCGCCCTCTTACAGTTAATGAAAAGAAAGAAATTAATTAATG 2639
Db 632 CysGlnGlnTyrValGlyProLeuThrValAsnGluLysArgLeuLysLeuIleMet 651
QY 2640 CCTGCTAGTTTATCTTAACCTTACTAAATATTTGCTTACAGCAAGGCAATTAACCG 2699
Db 652 ProAlaArgPheTyrProAsnLeuThrLysTyrLeuProLeuAspLysGlyIleLysPro 671
QY 2700 TATTATCTGAACATGCAAGTTAATCACTTACTTCAAACTAGGCATTATTTACATCTG 2759
Db 672 TyrTyrProGluHisAlaValAsnHisTyrPheLysThrArgHisTyrLeuHisThrLeu 691
QY 2760 TGGAGGCTGCACTCTATATAAGAGAAACTACAGCGCGCTCATTTTGTGGGTCA 2819
Db 692 TrpLysAlaGlyIleLeuTyrLysArgGluThrThrArgSerAlaSerPheCysGlySer 711
QY 2820 CCATATTTCTGGGACACAGACTACAGCATCGAGGTTGTGCTTCCAACTCCGACAGG 2879
Db 712 ProTyrSerTrpGluGlnGluLeuGlnHisGlyArgLeuValPheGlnThrSerGluArg 731
QY 2880 CATGGGAGCACTTGTGTTGCCAATCTCTCGGATCTTTTCCGATCAACCAAGTTGGA 2939
Db 732 HisGlyAspGluSerPheCysSerGlnSerSerGlyIleLeuSerArgSerProValGly 751
QY 2940 CCTGCGTTCCGAGCAACTCAAAATCCAGATTGGGACTTCAACCCCAACAGGATCA 2999
Db 752 ProCysIleArgSerGlnLeuLysGlnSerArgLeuGlyLeuGlnProGlnGlnGlySer 771
QY 3000 CTGGCCAGGCAATCAGGTAGGATGGGAGCATTCGGGCCAGGTTTACCCACACACA 3059
```

```
Db 772 MetAlaArgGlyLysSerGlyArgSerGlySerIleArgAlaArgValHisProThrThr 791
QY 3060 CGGCGCTCTTTGGGGGAGCCCTCAGCTCAGGCGATATTTCACACAGTCCGACGAGC 3119
Db 792 ArgArgSerPheGlyValGluProSerGlyThrArgHisIleAspAsnHisAlaSerSer 811
QY 3120 ACCTCTCTCTGCTCCACCAATCGGAGTCCAGGACAGCAGCTTACTCCCATCTCTCCAC 3179
Db 812 SerSerSerCysLeuHisGlnSerAlaValArgLysThrAlaTyrAlaHisLeuSerThr 831
QY 3180 TCTAAGACAGCATCTCTCAGGCCACGCGAGTGGAA 3215
Db 832 SerLysArgGlnSerSerSerGlyHisAlaValGlu 843

RESULT 2
Q67907_HPBVO
ID Q67907_HPBVO PRELIMINARY; PRT; 832 AA.
AC Q67907;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE DNA polymerase (Fragment).
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RC NUCLEOTIDE SEQUENCE.
RC STRAIN=Sub-type ayw;
RX MEDLINE=94079539; PubMed=8257295;
RA Preisler-Adams S., Schlayer M.J., Peters T., Hettler F., Gerok W.,
RA Rassenack J.;
RT "Sequence analysis of hepatitis B virus DNA in immunologically
RT negative infection.";
RL Arch. Virol. 133:385-396(1993).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; X72702; CAA51254.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000201; DNAPol_viral_N.
DR InterPro; IPR000477; RTase.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; RVT_1; 1.
DR ProDom; PD000814; DNAPol_viral_C; 1.
DR PROSITE; PSS00878; RT_POL; 1.
DR FT NON TER 1 1
DR FT NON TER 832 832
SQ SEQUENCE 832 AA; 93864 MW; EC6872097F04BAAF CRC64;

Alignment Scores:
Pred. No.: 8,42e-271 Length: 832
Score: 3745.50 Matches: 726
Percent Similarity: 72.16% Conservative: 49
Best Local Similarity: 67.60% Mismatches: 57
Query Match: 63.41% Indels: 243
DB: 2 Gaps: 2

US-10-761-006A-1 (1-3215) x Q67907_HPBVO (1-832)
QY 1 CTCACCAACATTTCCACCAAGCTCTGCTAGATCCAGGCTGAGGCGCTATATTTCTCTGC 60
Db 1 LeuHisAsnPheProProAsnSerAlaArgSerGlnGlyGluArgProValPheProCys 20
QY 61 TGGTGGCTCAGATTCCGGAAACAGTAACCCCTGTTCCGACTACTGCTCTCCCATATCGTC 120
Db 21 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 40
```

Qy	121	AATCTTTTCGAGGACTGGGGACCGTGCACCGAA	CATGGAGAACAACA	CATCAGGATTCC	180
Dd	41	AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisIleAeGliePro	:	:	60
Qy	181	AGGACCCCTCGTGTGTACAGGGGGTTTTCTCGTTGACAGAATCCTCAATAACC			240
Dd	61	ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr			80
Qy	241	GCAGAGTCTAGACTCTG- GTGGACTTCTCTCAATTTTCTAGGGGGAGCACCCCACGTGTTC			299
Dd	81	AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnHisAeGVal-Se	:	:	100
Qy	300	CTGGCCAAAATTGCACTGCCAACCTCCAATCACTCACCAACTCTTGTCCTCCAAATT			359
Dd	100	rTrpPrOlysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe			120
Qy	360	GTCCTGGCTACTCGCTGCATGTGTCGGGGCTTTTATCATATCTCTTCTCATCTGCTGC			419
Dd	120	uSerTrpLeuSerLeuAspValSerAlaAlaPheTyzHisLeuProLeuHisProAlaSe	:	:	140
Qy	420	TATGCTCATCTTCTTCTGTTCTTCTGCACTACCAAGTATGTGCCCGTTTGCTCT			479
Dd	140	rMetProHisLeuLeuValGlySerThrGlyLeuSerArgTyzValAlaAeGValSerSe	:	:	160
Qy	480	TACTTTCAGGAACATCAACACACGACGCGGGCCATCGACAGACTCGCACACTCCTGCTC			539
Dd	160	rAsnSerArgIlePheAsnHisGlnAeGlyThrMetGlnAsnLeuHisAspTyzCysSe	:	:	180
Qy	540	AAGGAAACTCTACTGTTTCCCTCTTGTGTGCTGTCACAAAACCTTCGGACGGAAACTGCACATT			599
Dd	180	rArgAsnLeuTyzValSerLeuLeuLeuLeuTyzGlnThrPheGlyArgLysLeuHisLe	:	:	200
Qy	600	GTATTCCTCCATCATCTCTGGGCTTTCGAGAGATTCCTATGGAGTGGGCTCAGTCC			659
Dd	200	uTyzSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr			220
Qy	660	GTTTCTCCTGGCTCAGTTTACTAGTGCATTTGTTCAGTGGTTCGTAGGCTTTCCCCCA			719
Dd	220	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValAeGValAePheProHi	:	:	240
Qy	720	CTGTTTCGCTTTTCAGTTATATGGATGATGTGTAATGGGGCGAAGTCTGTACAACTCT			779
Dd	240	sCysLeuAlaPheSerTyzMetAspAspValValLeuGlyAlaLysSerValGlnHisLe			260
Qy	780	TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGCTCTTGGGTATACATTTAAACCC			839
Dd	260	uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr	:	:	280
Qy	840	TAATAAAACAAACGTTGGGGCTACTCCCTTAACCTCATGGATATGTAATGGAAAGTTG			899
Dd	280	oAsnLysThrLysAeGrpGlyTyzSerLeuHisPheMetGlyTyzValIleGlySerTy	:	:	300
Qy	900	GGGTACTTTACCGCAGAAATATTGACTATAAACTCAAGCAATGTTTTCGAAAACTGCC			959
Dd	300	rGlySerLeuProGlnAspHisIleIleGlnLysIleGlyCysPheArgLysLeuPr	:	:	320
Qy	960	TGTAAATAGACCTATTGATTGGAAAGTATGTCAAAGAATTTGTGGCTCTTTTGGCTTTGC			1019
Dd	320	oIleAsnAeGrpIleAspTrpLysValCysGlnAeGrpIleValGlyLeuLeuGlyPheAl	:	:	340
Qy	1020	TGCCCTTTTACAAATGTGGCTATFCTGCTTGATGCCCTTTATATGCATGTATACAACT			1079
Dd	340	aAlaProPheThrGlnCysGlyTyzProAlaLeuMetProLeuTyzAlaCysIleGlnSe			360
Qy	1080	TAGCAGGCTTTCACCTTCTCGCAACTTACAGGCGCTTTCTGTGTAAACAATATCTGAA			1139
Dd	360	rLysGlnAlaPheThrPheSerProThrTyzLysAlaPheLeuCysLysGlnTyzLeuAs	:	:	380
Qy	1140	CTTTTACCCGCTTCCCGGCAACGGTCCGGTCTCTGCGAAGTGTGTTGCTGACGCAACCCC			1199
Dd	380	nLeuTyzProValAlaAeGlnAeGrpGlyLeuCysGlnVaIPheAlaAspAlaThrPr	:	:	400

Qy	1200	C	A	C	T	G	A	T	G	G	G	C	T	T	G	C	C	A	T	C	A	G	C	C	A	T	C	A	G	C	C	A	T	G	G	A	C	A	C	T	T	T	G	C	T	G	C	T	C	T	1259																																																																																																																																																																																																																																																																	
Db	400	o	T	r	G	i	y	T	r	G	i	y	L	e	u	V	a	M	e	t	G	l	y	H	i	s	G	i	n	A	r	g	M	e	t	A	r	g	G	i	T	r	P	h	e	L	e	u	A	l	a	P	r	o	L	e	420																																																																																																																																																																																																																																																											
Qy	1260	G	C	G	A	T	C	C	G	G	A	C	T	C	T	A	C	A	G	C	T	T	G	C	T	G	C	A	G	C	G	G	T	C	T	G	C	A	G	A	A	1319																																																																																																																																																																																																																																																																										
Db	420	u	P	r	o	L	i	e	H	i	s	T	r	A	l	a	G	l	u	L	e	u	a	l	a	C	y	S	p	h	e	A	i	a	r	g	S	e	r	A	r	g	S	e	r	G	l	y	A	l	a	s	440																																																																																																																																																																																																																																																															
Qy	1320	A	C	T	A	T	C	G	G	A	C	G	A	A	C	T	C	T	T	G	C	T	C	T	C	T	C	G	A	A	A	T	A	C	A	C	T	C	T	T	T	C	C	A	T	G	G	C	1379																																																																																																																																																																																																																																																																			
Db	440	n	i	L	e	L	e	G	i	y	T	r	A	s	p	A	s	n	S	e	r	V	a	L	l	e	u	S	e	r	A	r	g	L	y	S	e	r	T	r	S	e	r	P	h	e	P	r	o	T	r	P	h	e	460																																																																																																																																																																																																																																																													
Qy	1380	G	C	T	A	G	G	T	G	C	T	C	C	A	A	C	T	G	C	G	G	A	C	T	C	T	T	T	G	C	T	A	C	T	G	C	T	C	C	G	C	G	C	1439																																																																																																																																																																																																																																																																								
Db	460	u	L	e	U	G	i	y	C	y	s	A	i	a	a	s	n	T	r	i	L	e	u	a	r	g	G	i	y	T	r	S	e	r	P	h	e	V	a	L	r	o	S	e	r	A	l	480																																																																																																																																																																																																																																																																				
Qy	1440	G	C	T	G	A	A	T	C	C	G	C	G	A	C	C	C	G	T	C	T	C	G	G	G	C	C	G	T	T	T	G	G	G	C	T	C	T	A	C	C	G	T	C	T	T	C	A	1499																																																																																																																																																																																																																																																																			
Db	480	a	L	e	u	A	s	n	P	r	o	A	i	a	s	P	r	o	S	e	r	A	r	g	G	i	y	A	r	g	L	e	u	G	i	y	I	l	e	P	h	e	A	r	g	P	r	o	L	e	u	A	r	500																																																																																																																																																																																																																																																														
Qy	1500	T	C	T	G	C	G	T	T	C	C	G	C	C	A	C	C	G	G	G	C	C	T	C	T	T	T	A	G	C	G	G	T	C	T	C	C	C	G	T	A	T	G	T	C	1559																																																																																																																																																																																																																																																																						
Db	500	g	L	e	u	P	r	o	P	h	e	A	r	g	P	r	o	T	r	T	h	r	G	i	y	A	r	g	T	r	S	e	r	L	e	u	T	y	r	A	i	a	s	P	r	o	S	e	r	V	a	l	P	r	520																																																																																																																																																																																																																																																													
Qy	1560	T	T	C	T	C	A	T	C	T	G	C	A	T	T	C	G	T	T	C	C	A	C	T	T	C	A	C	C	T	C	T	G	C	A	C	T	G	C	A	T	G	G	A	C	C	A	C	1619																																																																																																																																																																																																																																																																			
Db	520	o	S	e	r	H	i	s	L	e	u	P	r	o	V	a	L	g	V	a	L	H	i	s	P	h	e	a	L	a	S	e	r	P	r	o	L	e	u	H	i	s	V	a	L	a	T	r	P	a	r	g	P	r	o	P	r	540																																																																																																																																																																																																																																																										
Qy	1620	G	T	G	A	A	C	G	C	A	G	C	G	T	T	G	C	C	A	A	G	G	T	T	A	T	A	T	A	T	A	A	G	A	G	A	C	T	T	T	G	A	C	T	C	T	C	A	G	A	1679																																																																																																																																																																																																																																																																	
Db	540	o	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

```
Db 541 -----Met--ProLeuSerTyrGlnHisPheArgLysLeu 551
Qy 2340 CTGTTGTTAGACGAGGAGGTCCTCTAGAGAGAACTCCCTCGCTCGCAGACGAA 2399
Db 552 LeuLeuLeuAspGluAlaGlyProLeuGluGluLeuProArgLeuAlaAspGlu 571
Qy 2400 GGTCTCAATCGCCGTCGAGAGATCTCAATCTCGGATCTCAACGTTAGTATTCCT 2459
Db 572 AspLeuAsnArgArgValAlaGluAspLeuAsnLeuGlyAsnLeuAsnValSerIlePro 591
Qy 2460 TGGACTCATAGGTGGGAACTTACTGGGCTTTATTCTTCTACTGTACCTGCTTTAAT 2519
Db 592 TrpThrHisLysValGlyAsnPheThrGlyLeuTyrSerSerThrValProValPheAsn 611
Qy 2520 CCCGAGTGGCAAAATCTCTCTTCATCATTCATTACAGAGACATTAATAATAGA 2579
Db 612 ProHisTrpLysThrProSerPheProAsnIleHisLeuHisGlnAspIleLeLysLys 631
Qy 2580 TGTCAACAATATGTGGCCCTCTTACAGTTAATGAAAAGAGAGATTAAATTAATTATG 2639
Db 632 CysGluGlnPheValGlyProLeuThrValAsnGluLysArgArgLeuGlnLeuIleMet 651
Qy 2640 CCTCTAGGTTTATCTTAACCTTACTAAATATTGGCTTTAGACAAAGGATTAAACCG 2699
Db 652 ProAlaArgPheTyrProAsnValThrLysTyrLeuProLeuAspLysGlyIleLysPro 671
Qy 2700 TATTATCTGAACATGCACTGATTAATCATTACTCAAACTAGGCATTTATACATCTCTG 2759
Db 672 TyrTyrProGluHisLeuValAsnHisTyrPheGlnThrArgHisTyrLeuHisThrLeu 691
Qy 2760 TGGAGGCTGCATCTTATATAAGAGAGAACTACAGCGCGCTCATTTTGGGTGCTCA 2819
Db 692 TrpLysAlaGlyIleLeuTyrLysArgGluThrThrArgSerAlaSerPheCysGlySer 711
Qy 2820 CCATATTCTTGGGAACAGACTACAGCATGGGAGTTGGTCTTCCAAACCTCGACAAAG 2879
Db 712 ProTyrSerTrpGluGlnLeuGln----- 720
Qy 2880 CATGGGAGCAATCTGCTGTCCCAATCTCTGGATTTCTTCCGATCAACGAGTTGGA 2939
Db 721 HisGlyAlaGluSerPheHisGlnGlnSerSerGlyIleLeuSerArgProProValGly 740
Qy 2940 CCCTGCGTTCGGAGCAACTCAAAATCCAGATGGGACTTCAACCCCAACAGGATCA 2999
Db 741 SerSerLeuGlnSerLysHisArgLysSerArgLeuGlyLeuGlnSerGlnGlnHis 760
Qy 3000 CTGGCCAGAGCAATCAGGTAGGAGTGGGAGCATTCGGGCGAGTTTCACCCCAACCA 3059
Db 761 LeuAlaArgArgGlnGlnGlyArgSerTrpSerIleArgAlaGlyIleHisProThrAla 780
Qy 3060 CGGCGGTCTTTTGGGGGGAGCCCTCAGGCTCAGGCGATATTGACAACAGTGCCAGCAGC 3119
Db 781 ArgArgProPheGlyValGluProSerGlySerGlyHisAsnThrAsnLeuAlaSerLys 800
Qy 3120 ACCTCTCTCTGCTCCCAATCGGAGTTCAGGAGACAGGCTACTCCCTCTCCACCC 3179
Db 801 SerAlaSerCysIleTyrGlnSerProValArgLysAlaAlaTyrProAlaValSerThr 820
Qy 3180 TCTAAGACACAGTCATCTCTCAGGCCACCCAGTGGAA 3215
Db 821 PheGluLysHisSerSerSerGlyHisAlaValGlu 832
RESULT 3
Q69590 HPBV0
ID Q69590 HPBV0 PRELIMINARY; PRT; 842 AA.
AC Q69590;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Polymerase (Fragment).
GN Name=P;
OS Hepatitis B virus.
```

```
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Genotype e;
RX MEDLINE=94120723; PubMed=8291231;
RA Norder H., Courouce A.M., Magnus L.O.;
RT "Complete genomes, phylogenetic relatedness, and structural proteins
of six strains of the hepatitis B virus, four of which represent two
new genotypes.";
RL Virology 198:489-503 (1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Genotype e;
RA Norder H.M.L.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
EMBL; X75657; CAA53339.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000201; DNAPol_viral_N.
DR InterPro; IPR000477; RTase.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; RVT_1; 1.
DR ProDom; PD000814; DNAPol_viral_C; 1.
DR PROSITE; PS00878; RT_BOL; 1. Potential.
FT CHAIN <1> >842
FT NON TER 1
FT NON TER 842
FT SEQUENCE 842 AA; 94812 MW; 45C3CD71CF5A4PE7 CRC64;
SQ
Alignment Scores:
Pred. No.: 6,51e-267 Length: 842
Score: 3693.50 Matches: 719
Percent Similarity: 71.85% Conservative: 52
Best Local Similarity: 67.01% Mismatches: 70
Query Match: 62.53% Indels: 233
DB: 2
US-10-761-006A-1 (1-3215) x Q69590_HPBV0 (1-842)
Qy 4 CACAACATTCCACCAAGCTCTGCTAGATCCAGGGTGAGGGGCTATATTTCTCTGCTGG 63
Db 2 HisAsnIleProProSerSerAlaGlySerGlnSerLysArgProValPheSerCysTrp 21
Qy 64 TGGCTCCAGTTTCGGAAACAGTAACCCCTGTTCCGACTACTGCTCTCCCATATCGTCAAT 123
Db 22 TrpLeuGlnPheArgAsnSerGluProCysSerAspTyrCysLeuThrHisLeuValAsn 41
Qy 124 CTTCTCCAGGACTGGGACCCCTGCACCGAATCATGAGAACACATCAGGATTCCTAGG 183
Db 42 LeuLeuGluAspTrpGlyProCysThrGluHisGlyHisHisIleArgIleProArg 61
Qy 184 ACCCTCTGCTGTTTACAGCGGGGTTTTCTCTGTTGACAGAAATCTCCACAAATACCCCA 243
Db 62 ThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThrAla 81
Qy 244 GAGTCTAGACTCTG-GTGGACTTCTCAATTTTCTAGGGGGAGACCCACGTTCTCTG 302
Db 82 GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerArgVal-SerTr 101
Qy 303 GCCAAATTCGACGTCCCACTCCCACTCACTCACCACCTTGTCTCTCCAAATTTGTC 362
Db 101 pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe 121
Qy 363 CTGGCTATCGTGGATGTCTGCGGGCGTTTTTATCATATTCCTCTTCATCTCTGCTGCTAT 422
```

|||||
121 rTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAlaIle 141
|||||
423 GCCTCATCTTCTGTTGGTCTTCTGAGACACCAAGGTATGTTGCCGTTTGCTCTAC 482
|||||
141 tProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSerAs 161
|||||
483 TTCCAGGACATCAACACCAAGCAGCGGCCATGCAAGACCTCAGACTCCCTGCTCAAG 542
|||||
161 nSerArgIleAAsnHisGlnTyrGlyThrLeuProAsnLeuHisAspSerCysSerAr 181
|||||
543 GAAACTCTAGCTTCCCTCTGTTGCTGTACAAACCTTCGGACGGAAACTGACACTTGTA 602
|||||
181 gAsnLeuTyrValSerLeuMetLeuLeuPheTyrHisPheGlyArgLeuHisLeuTyr 201
|||||
603 TTCCCATCCCATCATCTCGGCTTTCGCAAGATTCCTATGGGAGTGGGCTCAGTCCGTT 662
|||||
201 rSerHisProIleMetGlyPheArgIleProMetGlyValGlyLeuSerProPh 221
|||||
663 TCTCCTGGCTCAGTTTACTAGTCCCATTTGTTTCAGTGGTTCGTAGGCTTCCCCCAGT 722
|||||
221 eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCy 241
|||||
723 TTTCGGCTTTCAGTTATATGATGATGTTGTTATGGGGCGAAGTCTGTACAACTCTGA 782
|||||
241 sLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLeuGln 261
|||||
783 GTCCCTTTTACCTTATACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCTAA 842
|||||
261 uSerLeuTyrThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAs 281
|||||
843 TAAACCAACGTTGGGCTACTCCCTTAACCTTCATGGGATATGTAATTGGAAGTTGGG 902
|||||
281 nLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrpGln 301
|||||
903 TACTTTACCGCAGCAATATGCTATAAAGCTCAAGCAATGTTTTCGAAACCTGCTGT 962
|||||
301 ySerLeuProGlnGluHisIleGlnIleGlnLysLysAspCysPheArgLysLeuProVa 321
|||||
963 AAATAGACTATTGATGAAAGTATGTCAAGAATTGTGGGTCTTTTGGGCTTCTGCTGC 1022
|||||
321 lAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl 341
|||||
1023 CCTTTTACACATGTGGCTATCTGCTGCTGATGCTTATATGATGATGATACAACTAA 1082
|||||
341 aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysThrGlnSerLys 361
|||||
1083 GCAGGCTTTCACCTTCTCGCCAACTTACAAGGCTTCTGTGTAAACAATATCTGAACCT 1142
|||||
361 sGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsnLe 381
|||||
1143 TTACCCCGTTGCCGCAACGGTCCGCTCTGTCCAAAGTGTGTTGTGACGCAACCCCAAC 1202
|||||
381 uTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrProTh 401
|||||
1203 TGGATGGGCTTGGCATAGGCCATCAGCGCATGGCTGGAACCTTCTGCTCTCTGCTGCC 1262
|||||
401 rGlyTrpGlyLeuAlaIleGlyIleGlnArgMetArgGlyThrPheValAlaProLeuPr 421
|||||
1263 GATCCATCTATCGCGAACTCTACAGCTTGTGTTGCTCGCAGCGGCTCTGGAGCAAACT 1322
|||||
421 orleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLysLe 441
|||||
1323 TATCGGAACCGCAACTCTGTGTTGCTCTCTCGGAATATACCTCTCTTTCATGCTGCT 1382
|||||
441 uIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeuLe 461
|||||
1383 AGGCTGTGTCACACTCGATCTCGCGGAGCTCTTGTCTAGCTCCGCTCGGCGCT 1442
|||||
461 uGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLe 481
|||||
1443 GAATCCCGGAGCAGCCCGTCTCGGGCGGCTTGGGGCTCTACCGTCCCTCTTCTCATCT 1502
|||||

Db 481 uAsnProAlaAspAspProSerArgGlyArgLeuGlyIleTyrArgProLeuLeuArgLe 501
QY 1503 GCGGTTCCGGCGACCAACCGGGCGCACTCTCTTTACGGGCTCTCCCGTATGTGCTTC 1562
|||||
Db 501 uProPheGlnProThrGlyArgThrSerLeuTyrAlaValSerProSerValProSe 521
|||||
QY 1563 TCATCTGCGGACCGGTGTCACCTTCGCTTCACCTCTGACGCTGCATGGAGACCCGCTG 1622
|||||
Db 521 rHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro-- 540
|||||
QY 1623 AACGCAACCGCAGTCTTGCCCAAGGTCTTATATAGAGGACTCTTGGAATCTCAGCAATG 1682
|||||
Db 540 ----- 540
QY 1683 TCAACGACCGACTTGAGGCATCTTCAAAGACTGTGTGTTTAAAGACTGGGAGGAGTTG 1742
|||||
Db 540 ----- 540
QY 1743 GGGAGGAGATTAGGTTAAAGATTATGTACTAGGAGGCTGTAGGCATAAATTGCTGT 1802
|||||
Db 540 ----- 540
QY 1803 TCACCAGCACCATGCAACTTTTCTCCTCTGCTTAATCATCTCATGTTCATGTCCTACTG 1862
|||||
Db 540 ----- 540
QY 1863 TTCAAGCCTCAAGCTGTGCTTGGGTGGCTTGGGACATGGACATGACCCGTATAAAG 1922
|||||
Db 540 ----- 540
QY 1923 AATTTGGACATCTGCTGAGTTACTCTCTTTTGGCTTCTGACTTCTTCCGCTATTTC 1982
|||||
Db 540 ----- 540
QY 1983 GAGATCTCTCGACACCGCTCTGCTCTGTATCGGAGGCTTAGAGTCTCCGGAACATT 2042
|||||
Db 540 ----- 540
QY 2043 GTTCGCTCACCATACAGCCTCAGGCAAGCTATTTTGTGTTGGGTGAGTTGATGAATC 2102
|||||
Db 540 ----- 540
QY 2103 TGGCCACCTGGGTGGGAAGTAATTTGGAAGATCCAGCATCCAGGGAATTAGTAGTCAGCT 2162
|||||
Db 540 ----- 540
QY 2163 ATGTCAACCTTAATATGGGCTTAAAACTCAGACAAATATTTGTTTTCACATTTCTCTGTC 2222
|||||
Db 540 ----- 540
QY 2223 TTACTTTTGGAAAGAAACTGTTCTTGAGTACTTGTGTATCTTTTGGAGTGTGGATTCGCA 2282
|||||
Db 540 ----- 540
QY 2283 CTCCTACCGCTTACAGACCACCAATGCCCCCTTATCTTAAACACTTCGGAACACTACTG 2342
|||||
Db 541 -----Met--ProLeuSerTyrGlnHisPheArgGlnLeu 552
|||||
QY 2343 TTGTTTACGACGACGAGCGAGTCCCTCTAGAAAGAACTCCCTCGCTCGCAGACGGAAGGT 2402
|||||
Db 553 LeuLeuAspGluGluAlaGlyProLeuGluGluLeuProArgLeuAlaAspGluAsp 572
|||||
QY 2403 CTCATCGCGGCTGCGAGAACTCTCAATCTCGGGAATCTCAACGTTAGTATTCCTTGG 2462
|||||
Db 573 LeuAsnArgArgValAlaGluAspLeuAsnLeuGlnLeuProAsnValSerIleProTrp 592
|||||
QY 2463 ACTCATAGCTGGAAACTTTTACTGGGCTTATCTTCTACTGTACTCTCTTCTTAAATCCC 2522
|||||
Db 593 ThrHisLysValGlyAsnPheThrGlyLeuTyrSerSerThrIleProValPheAsnPro 612
|||||
QY 2523 GAGTGGCAAAATTCCTTCTCTCTTCTCACATTCATTTTCAAGAGGACATTTAATAGATGT 2582
|||||
Db 613 AsnTrpLysThrProSerPheProAspIleHisLeuHisGlnAspIleIleAsnLysCys 632
|||||

```
QY 2583 CAACAATATGTGGCCCTCTTACAGTTAAATGAAAAGAGATTAAATTAATTATGCTT 2642
Db 633 GluGlnPheValGlyProLeuThrValAsnGluLysArgArgLeuAsnLeuValMetPro 652
QY 2643 GCTAGGTTTATCTTAACTTACCTTAAATATTTGCCCTTAGACAAAGGCATTAACCGTAT 2702
Db 653 AlaArgPhePheProIleSerThrLysTyLeuProLeuGluLysGlyIleLysProTyr 672
QY 2703 TATCTGGAACATGCAAGTAAATCAATCTTCAAACTAGGATTTATTTACATCTCTGTGG 2762
Db 673 TyrProAspAsnValValAsnHisTyPheGlnThrArgHisTyLeuHisThrLeuTrp 692
QY 2763 AAGCTGGCAATCTATATAGAGAACTACAGCAGCGCCCTCATTTCTGGTGGTCAACA 2822
Db 693 LysAlaGlyIleLeuTyLysArgGluThrArgSerAlaSerPheCysGlySerPro 712
QY 2823 TATCTTGGGAACAAGAGCTTACAGCATGGGAGGTTGGTCTTCCAAACCTCGACAAGGCAT 2882
Db 713 TySerTrpGluGlnGluLeuHisGlyAlaPheLeu--AspGlyProSerArgMet 731
QY 2883 GGGAGCAATCTTGCTCTTCCCAATCTCTGGGATTTCTTCCGATCACCAGTTGGACCC 2942
Db 732 GlyGluGluTyPheHisHisGlnSerSerGlyIlePheSerArgProProValGlySer 751
QY 2943 TGGCTCGGAGCCAACTCAACAATCCAGATTGGGATTTCAACCCCAACAGGATCACTG 3002
Db 752 SerIleGlnSerLysHisGlnLysSerArgLeuGlyProGlnSerGlnArgProLeu 771
QY 3003 GCCAGAGCAATCAGGTAGGAGTGGGAGCATTCGGCCAGGGTTCAACCCACACACGG 3062
Db 772 AspArgSerGlnGlnGlyArgSerGlySerIleArgAlaTrpValHisSerProThrArg 791
QY 3063 CGGTCTTTTGGGGGGAGCCCTCAGGCTCAGGCGATATTGACAACTGCGCAGCAGCAC 3122
Db 792 ArgProPheGlyValGluProSerGlySerArgHisAlaLysAsnIleAlaSerArgSer 811
QY 3123 TCCTCTGCTCCCAATCGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 3182
Db 812 AlaSerCysLeuHisGlnSerAlaValArgLysAlaLysAlaTyProAsnHisSerThrPhe 831
QY 3183 AAGAGCAGCATCTCCTCAGGCGCAGCAGTGGAA 3215
Db 832 GluArgHisSerSerSerGlyHisAlaValGlu 842

RESULT 4
ID Q7TDR5_HPBVO PRELIMINARY; PRT; 843 AA.
AC Q7TDR5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymerase protein.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Liu C.-J., Chen P.-J., Lai M.-Y., Kao J.-H., Chen D.-S.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY167095; AAO41311.1; -; Genomic DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006278; F:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000201; DNAPol_viral_N.
DR InterPro; IPR000477; RVtse.
```

```
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; RVT_1; 2.
DR ProDom; PD000814; DNAPol_viral_C; 1.
SQ SEQUENCE 843 AA; 94471 MW; A40B9C18D84B4EA5 CRC64;

Alignment Scores: 6.32e-202 Length: 843
Pred. No.: 2824.00 Matches: 530
Score: 98.52% Conservative: 3
Percent Similarity: 97.97% Mismatches: 7
Best Local Similarity: 47.81% Indels: 2
Query Match: 2 Gaps: 0

US-10-761-006a-1 (1-3215) x Q7TDR5_HPBVO (1-843)

QY 1 CTCACACATTTCCACCAAGCTCTGTAGTACCAGGGTGGAGGGCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProSerSerThrArgSerGlnSerGlnGlyProIlePheSerCys 323
QY 61 TGGTGGCTCCAGTTCCGGAACAGTAAACCTGTTCGACTACTGCTCTCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgHisSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTCGAGACTGGGACCTCGCACCGAATCGAGAACACACATCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGACCCCTCTCGTTTACAGCGGGGTTTTCTCGTTGACAGATCTCTCAATATACC 240
Db 364 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGTCTAGACTCG-GTGGACTTCTCTCAATTTTCTAGGGGAGACCCACGCTGTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgLysSerThrHisVal-Se 403
QY 300 CTGGCCAAAATTCGAGTCCCACTCCCAATCCTCACCACCTCTTGCTCTCCAATTT 359
Db 403 rTTPProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY 360 GTCTGCTATCGCTGGATGTCTGGCGGTTTTTATCATATTCTCTTCTCTCTCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyHisIleProLeuHisProAlaAl 443
QY 420 TATGCTCATCTCTCTGTTCTCTGCTGACTACCAAGGTATGTTGCCCTTTGCTCTCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACACACGACGCGGGCCATGCAAGACCTGACGACTCTGCTGC 539
Db 463 rThrSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAACTCTACGTTTCCCTCTTGTGCTGTGTACAAAACCTTCGACGGAACCTGCACTT 599
Db 483 rArgAsnLeuTyValSerLeuLeuLeuTyLysThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTCATCCATCCATCTCTGGCTTTCGCAAGATTCCTATGGAGTGGGCGCTCAGTCC 659
Db 503 uTySerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
QY 660 GTTCTCTCTGCTCAGTTTACTAGTGCATTTGTTTCTAGTGTGTTCTGAGGGCTTTCCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
QY 720 CTGTTTGGCTTTTCAAGTTATATGAGATGATGTTGGTATTTGGTGGCGGAGTCTGTACA 779
Db 543 sCysLeuAlaPheSerTyMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACCTCTATACCAATTTTCTTTTCTTTTGGTGTATACATTTTAAACCC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACTTTCATGGGATATGTAATTGGAAGTTG 899
```

Db 583 oAnlystThLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValleGlySerTr 603
QY 900 GGGTACTTTTACCGCAGGAACATATTGTACTAAACCTCAAGCAATGTTTTCGAAAACTGCC 959
Db 603 pGlyThrLeuProGlnGluHisIleValleuLysLeuLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACCTATTGATTGGAAAGATGTCAAAGAAATGTGGGCTTTTGGGCTTTGCG 1019
Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuGlyPheAl 643
QY 1020 TGCCCTCTTTACACAATGTGGCTATCTGCTCTGTGATGCCCTTTATATGCGATGTATACAATC 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAGCAGGCTTTTCACTTTCTCGCAACTTACAGGCTTTTCTGTGTAAACAATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 683
QY 1140 CTTTACCCGTTGCGCGGCAAGGTCGGTCTCTGCCAAGTGTGTGCTGACGCAACCC 1199
Db 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
QY 1200 CACTGTGATGGGCTTGGCCATAGCCATCAGCGCATGGCTGGAACTTTCTGGCTCTCT 1259
Db 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 723
QY 1260 GCCGATCCCATCTGCGAACTCTTAGCAGCTTGTGCTCGAGCGGCTCTGGAGCAAA 1319
Db 723 uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 743
QY 1320 ACTTATCGGAACCGACAACCTGTTGCTCTCTCGGAATATACACTCTCTTCCATGGCT 1379
Db 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
QY 1380 GCTAGGCTGTGTCGCAACTGGATCTGCGCGGAGCTCTTGTCTGTACGTCCGTCGCGC 1439
Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
QY 1440 GCTGAATCCCGCGAGACGCCGCTCTCGGGCCGCTTGGGGCTCTACCGTCCCTCTTCA 1499
Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
QY 1500 TGTGCGGTTTCCGCGCGACACGCGGCGCACCTCTCTTTAGCGGTCTCCCGTATGTGCC 1559
Db 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
QY 1560 TTCTCATCTGCGGACCGGTGTGCATTCGCTTCCACTCTGCACTGCGCATGGAGACCAC 1619
Db 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaIleTrpArgProPr 843
QY 1620 G 1620
Db 843 o 843

RESULT 5

Q4FD83 HPBV0
ID Q4FD83 HPBV0 PRELIMINARY; PRT; 843 AA.
AC Q4FD83;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Polymerase.
DE Name=P;
GN Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthonepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HK343;
RX PubMed=15897987; DOI=10.1086/430324;
RA Chan H.L., Tsui S.K., Tse C.H., Ng E.Y., Au T.C., Yuen L.,

RA Bartholomeusz A., Leung K.S., Lee K.H., Locarnini S., Sung J.J.;
RT "Epidemiological and virological characteristics of 2 subgroups of
RT hepatitis B virus genotype C";
RL J. Infect. Dis. 191:2022-2032(2005).
DR EMBL; DQ089793; AA05280.1; -; Genomic DNA.
SQ SEQUENCE 843 AA; 94494 MW; BB50233DPEBA4E29 CRC64;

Alignment Scores:
Pred. No.: 8,92e-202 Length: 843
Score: 2822.00 Matches: 530
Percent Similarity: 98.52% Conservative: 3
Best Local Similarity: 97.97% Mismatches: 7
Query Match: 47.77% Indels: 2
DB: Gaps: 0

US-10-761-006A-1 (1-3215) x Q4FD83_HP BV0 (1-843)

QY 1 CTCACAACATTTCCACCAAGCTCTGTAGATCCCAAGGTGAGGGCGCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
QY 61 TGGTGGCTCCAGTTCGGGAACAGTAAACCTCTTCCGACTACTGCTCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTCAGAGACTGGGGACCTGACCGAACATGAGAGAACACACATCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGAGCCCTGCTCGTGTACAGCGGGGTTTTCTCGTTGACAAAGATCCTCACAAATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGTCTAGACTCTG-GTGACTTCTCTCAATTTTCTAGGGGGAGCACCCACGTGTT 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY 300 CTGGCCAAAATTCGAGTCCCCAACCTCCAATCCTACCAACCTCTTGTCTCTCCAAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY 360 GTCCTGGCTATCGCTGGATGCTCTGCGCGGTTTTATCATATTTCTCTTTCATCTCTGTC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
QY 420 TATGCTCATCTCTTGTGTGTTCTTCTGAGCTACCAAGGTATGTTGCCCGTTGTCTCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACCAACAGCAGCGGGCCATGCAAGACCTGACAGCTCTCTGCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAACCTACGTTTCCCTCTGTTGCTGTACAAAACCTTCGGACGGAACCTGCACATT 599
Db 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTCCCATCCCATCATCTCTGGGCTTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
Db 503 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
QY 660 GTTCTCTCGGCTCAGTTTACCTAGTGCCATTTGTCAGTGGTTCGTAGGGCTTTTCCCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
QY 720 CNGTTTGGCTTTTCAGTTATATGATGTGTATTGGGGCGAAGCTGTGTACAACATCT 779
Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACTCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583

QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAATCTCATGGATATGTAATGGAAGTTG 899
DB 583 oAenLysThrLysArgTrpGlyTyrSerLeuAnpHeMetGlyTyrValIleGlySerTr 603
QY 900 GGTGACTTTTACCAGGAAACATATGTTACTAAACAACTCAAGCAATGTTTTCGAAAATCGCC 959
DB 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACCTATTGATTGGAAAGATATGTTCAAGAAATGTTGGGCTTTTGGCTTTC 1019
DB 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGGCCCTTTTACACAATGTCCTATCTGCTGATGCTCTGATGCTCTTATATGATGATATCAATC 1079
DB 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAGCAGGCTTTTCACTTTCTCGCAACTTACAAGGCTTTCTGTGTAAACAATATCTGAA 1139
DB 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuLysGlnTyrLeuAs 683
QY 1140 CTTTACCCGTTGCGCGCAACGGTCCGCTCTCTGCAAGTGTGTTGCTGACGCAACCCC 1199
DB 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
QY 1200 CACTGGATGGGCTTGGCCATAGCCATCAGCGCATCGCTGGAACCTTTCTGGCTCTCT 1259
DB 703 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 723
QY 1260 GCCGATCCATCTCGGAACTCTCTAGCAGCTGTGTTTGTGCGACGCGTCTGAGCAAAA 1319
DB 723 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLy 743
QY 1320 ACTTATCGAACCAGCAACTCTGTGTCCTCTCTCGGAATAATACCTCTCTTCCATGCT 1379
DB 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
QY 1380 GCTAGGTGTGCTGCCAAGTCTGCTGCGGACGCTCTTCTGCTACGTCGCGTCGCG 1439
DB 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
QY 1440 GCTCAATCCGCGGACGACCTCTCGGGGCGGTTTGGGCTCTACCGTCTCCCTTCTTCA 1499
DB 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
QY 1500 TCTCCGTTCCGCGGACGACGCGGCGGACCTCTCTTCTTACCGGCTCTCCCGTATGTC 1559
DB 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
QY 1560 TTCTCATCTGCGGACGCTGTGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1619
DB 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
QY 1620 G 1620
DB 843 o 843
RESULT 6
ID Q69026_HPBVO PRELIMINARY; PRT; 540 AA.
AC Q69026;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE P gene product (Fragment).
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AYT;
RX MEDLINE=87059755; PubMed=3783127;
RA Okamoto H., Imai M., Shimozaki M., Hoshi Y., Iizuka H., Gotanda T.,

RA Tsuda F., Miyakawa Y., Mayumi M.;
RT "Nucleotide Sequence of a Cloned Hepatitis B Virus Genome, Subtype
RL avr:Comparison with Genomes of the Other Three Subtypes."; J.
CC J. Gen. Virol. 67:2305-2314 (1986).
DB -1- SIMILARITY: Contains 1 reverse transcriptase domain.
EMBL; X04615; CAA28286.1; -; Genomic_DNA.
DR GO: 0003677; F:DNA binding; IEA.
DR GO: 0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO: 0004523; F:ribonuclease H activity; IEA.
DR GO: 0003723; F:RNA binding; IEA.
DR GO: 0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: 0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00078; RVT_1; 1.
DR ProDom; PD000814; DNAPol_viral_C; 1.
DR PROSITE; PS00878; RT_POL; 1.
FT NON_TER 1
SQ SEQUENCE 540 AA; 60365 MW; 80446CA7C80F998A CRC64;
Alignment Scores: Length: 540
Pred. No.: 6 02e-201 Matches: 527
Score: 2810.00
Percent Similarity: 98.34%
Best Local Similarity: 97.41%
Query Match: 47.57%
DB: 2
US-10-761-006A-1 (1-3215) x Q69026_HPBVO (1-540)
QY 1 CTCACAACTTCCACCAAGCTCTGCTAGATCCAGGGTCCAGGGCTATATTTCTCTGC 60
DB 1 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 20
QY 61 TGTGGCTCCAGTTCGGAAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTC 120
DB 21 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 40
QY 121 AATCTTCTCAGGACTGGGACCTCGCACCGAATCATGAGAACACACATCAGGATTCCT 180
DB 41 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 60
QY 181 AGGACCTCTGCTGTTTACAGCGGGGTTTTCTCGTTGACAAAGATCTCTCAATACC 240
DB 61 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 80
QY 241 GCAGAGTCTAGACTGTG-GTGGACTTCTCTCAATTTCTAGGGGAGCACCACGTGTTTC 299
DB 81 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 100
QY 300 CTGGCCCAAAATTCGACGTCCTCCCACTCCCAATCACTCAACCACTCTTGTCTTCCAATT 359
DB 100 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 120
QY 360 GTCCTGCTATCGTGGATGTCTGCGGGGTTTTTATCATATTTCTCTTCTCTCTGCTGC 419
DB 120 uSerTrpLeuSerLeuAspValSerAlaPheTyrHisIleProLeuHisProAlaAl 140
QY 420 TATGCCTCATCTTCTGTTGTTCTTCTGGAATCAAGGATGTTGTCCTGTTGTCCTC 479
DB 140 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 160
QY 480 TACTTCCAGGACATCAACACGACGCGGGCCCATGCAAGCACTTCCGACGGAACCTCCT 539
DB 160 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 180
QY 540 AAGGAACTCTACGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 599
DB 180 rArgAsnLeuTyrValSerLeuLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHis 200
QY 600 GTATTTCATCCATCATCTCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659

Db 200 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 220
QY 660 GTTTCCTCGGCTCAGTTACTAGTCCATTTGTTCAAGTGTTCGTAGGGCTTTCCCA 719
Db 220 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValArgAlaPheProHi 240
QY 720 CTGTTTGGCTTTTCTAGTATATAGTATGATGTGGTATTTGGGGCCCAAGTCTGTACAAATCT 779
Db 240 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 260
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGTTCTTTGGGTATACATTTAAACCC 839
Db 260 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 280
QY 840 TATAAACAACCGTTGGGCTACTCCCTTACTTCACTGCGATATGTAATTGGAAGTTG 899
Db 280 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 300
QY 900 GGTGTACTTTACCGCAGCAATATTTGTAATAAATCAAGCAATGTTTTCGAAAACCTGCC 959
Db 300 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 320
QY 960 TGTAAATAGACCTATTGATTGGAAGTATGTCAAAGAAITTTGGGCTTTTGGCCTTTGC 1019
Db 320 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 340
QY 1020 TGCCCTTTTACATATGGGCTATCTGCTGTATGCTTATATGATGATGATATACAATC 1079
Db 340 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 360
QY 1080 TAAGCAGGCTTTTCTGCGCAACTTCAAGGCTTTCTGCTGTAACAATATCTGAA 1139
Db 360 rLysGlnAlaPheThrPheSerProThrLysAlaPheLeuLysGlnTrpLysLeuAs 380
QY 1140 CCTTTACCCCGTTGCCGCAACGTCGCTCTCTGCCAAGTGTGTTGCTGACGCAACCCC 1199
Db 380 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 400
QY 1200 CACTGATGGGCTTGGCCATAGGCCATAGCCATGCGCATGGCTGGAACTTTCTGGCTCTCT 1259
Db 400 oThrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 420
QY 1260 GCGATCCCATCTAGCGAACTCTAGCAGCTTGTGTTGCTCGCAGCGGCTGGAGCAAA 1319
Db 420 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLys 440
QY 1320 ACTTATCGGAACGCAACTCTGTTCTCTCTCGGAAATACACCTCTCTTCCATGGCT 1379
Db 440 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 460
QY 1380 GCTAGGCTGTGCGCAACTGGAATCTGCGCGGAGCGTCTTGTCTAGTCCCGTCGCGC 1439
Db 460 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 480
QY 1440 GCTGAATCCCGGAGCAGCCGCTCGGGCGCTTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db 480 aLeuAsnProAlaAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 500
QY 1500 TCTGCGGTTCCCGCCACACCGGGCGCACCTCTCTTTACCGGCTCTCCCGCTATGTC 1559
Db 500 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 520
QY 1560 TTCTCATCTGCGGACCGGTGTGCACTTGTTCACCTCTGACGTCGATGGAGACCAACC 1619
Db 520 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 540
QY 1620 G 1620
Db 540 o 540

RESULT 7
Q9QN49 HPBV0
ID Q9QN49_HPBV0 PRELIMINARY; PRT; 843 AA.

AC Q9QN49; 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 26, Last annotation update)
DE Polymerase.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=genotype C;
RX MEDLINE=20528425; PubMed=11074476;
RX DOI=10.1002/1096-9071(200012)62:4<4471::AID-JMV12>3.3.CO;2-O;
RA Rokuhara A., Tanaka E., Yagi S., Mizokami M., Hashikura Y.,
RA Kawasaki S., Kiyosawa K.;
RT "De novo infection of hepatitis B virus in patients with orthotopic
RT liver transplantation: analysis by determining complete sequence of
RT the genome.";
RL J. Med. Virol. 62:471-478(2000).
DR EMBL; AB025815; BAA85065.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000201; DNAPol_viral_N.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; RVT 1; 2.
DR ProDom; PD000814; DNAPol_viral_C; 1.
SQ SEQUENCE 843 AA; 94462 MW; 9D741241062DEAAE CRC64;

Alignment Scores:
Pred. No.: 9,92e-201 Length: 843
Score: 2808.00 Matches: 527
Percent Similarity: 98.52% Conservative: 6
Best Local Similarity: 97.41% Mismatches: 7
Query Match: 47.54% Indels: 2
DB: Gaps: 0

US-10-761-006A-1 (1-3215) x Q9QN49_HPBV0 (1-843)
QY 1 CTCCCAACATTCCACCAGCTCTCTAGATCCAGGGTGGGGCTATATTTCTCTGC 60
Db 304 LeuHisHisIleProProSerSerAlaArgSerGlnSerGlyProIlePheSerCys 323
QY 61 TGTGTGCTCCAGTTCGGGAACAGTAAACCTCTCTCGACTACTGCTCTCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTCAGGACTGGGGACCTCGACCAACATGGAGAACACAACTCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTCGTGTTCACGGCGGGTTCCTCGTTGACAAAGTATCCTCACATACC 240
Db 364 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 CGAGACTCTAGACTCTG-GTGGACTTCTCTCAATTTCTTAGGGGGAGCACCCACGTGTT 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY 300 CTGGCCAAAATTGCGAGTCCCCCAACTCAATCACTCACCACCTTCTCTCTCCCAATT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY 360 GTCCTGGCTATCGCTGGATGTCTCTGCGCGTGTTCATCATATTTCTCTCTCTCTCTCT 419

Db 423 userTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
Qy 420 TATGCCCTCATCTCTTCTGTTCTCTGAGTACCAAGTATGTCGCCGTTGTCCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
Qy 480 TACTTCCAGGACATCAACACGACGACGGCCATGCAAGACCTGCAACGACTCCCTGCTC 539
Db 463 rTyrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
Qy 540 AAGGAAACTCTACGTTTCCCTCTTGTGTCTGTATCAAAACCTTTCGACGGAACCTGCACTT 599
Db 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
Qy 600 GTATTCCTCCATCCATCATCTGGCTTTCGCAAGATTCCTATGGAGTGGCCCTCAGTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy 660 GTTTCCTCTGCTCAGTTTACTAGTGCATTTGTTCTGAGTTCGTTAGGCTTTCCCCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 543
Qy 720 CTGTTTCTGCTTTCAGTTATATGATGATGTGTATTTGGGGCGAAGTCTGTACAACTCT 779
Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
Qy 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTCTTTGTTGGGTATACATTTAAACCC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
Qy 840 TAATAAACCAACGTTGGGCTACTCCCTTAACTCATGCGATATGTAATGCAAGTTG 899
Db 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
Qy 900 GGGTACTTTTACCGAGGAACATATTGTACTAACTCAAGCAATGTTTTCGAAACTGCC 959
Db 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
Qy 960 TGTAAATAGACTTATGATGGAAGTATGTCAAGAACTGTGGGCTTTTGGGCTTTGC 1019
Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Qy 1020 TGCCCTTTTACAAATGTGCTATCTGCTGCTGCTGCTTATGATGATGATGATGATGATC 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
Qy 1080 TAACGAGCTTTCATCTTCTCGCAACTTACAGGCTTTCTGTGTAAACAATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
Qy 1140 CTTTACCCCGTGGCCGCAACCGTCCGCTCTCTGCCAAGTGTGTTGTCGACGCAACCCC 1199
Db 683 sLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
Qy 1200 CACTGATGGGCTTGGCCATAGCCCATCAGCGCATGGCTGGAACCTTCTGGCTCTCTCT 1259
Db 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 723
Qy 1260 GCCGATCCATCTCGGAACCTCTAGCAGCTGTTTCTGTCAGCCGCTCTGGAGCAAA 1319
Db 723 uProfileHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLys 743
Qy 1320 ACTTATCGAACCGCAACTCTGTTGTCTCTCTCGGAATACACCTCTTTTCCATGCTCT 1379
Db 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTiple 763
Qy 1380 GCTAGGGTGTGTGCAACTGATCCTGCGGGAGCTCTTTGCTAGCTCCCTCGGC 1439
Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
Qy 1440 GCTGAATCCGCGGAGGACCGCTCTCGGGCGGTTTGGGCTCTACCGTCCCTCTTCTCA 1499

783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
Qy 1500 TCTCGCGTTTCCGCGCCACACACGCGCGCCTCTCTTTACGCGGTCTCCCGGTATGTGCC 1559
Db 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
Qy 1560 TTCTCATCTCGCGGACCGTGTGCATCTTCCTTCACCTTCGACGTCGATGGAGACACC 1619
Db 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
Qy 1620 G 1620
Db 843 o 843
RESULT 8
DPOL_HPBVR STANDARD; PRT; 843 AA.
ID AC P03157;
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN Name=P;
OS Hepatitis B virus (subtype adr).
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=106820;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RA Ono Y., Onda H., Sasada R., Igarashi K., Sugino Y., Nishioka K.;
RT "The complete nucleotide sequences of the cloned hepatitis B virus
DNA; subtype adr and adr."
RL Nucleic Acids Res. 11:1747-1757(1983).
CC -1- CATALYTIC ACTIVITY: Deoxynucleoside triphosphate + DNA (n) =
diphosphate + DNA(n+1).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.

This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

EMBL; V00867; -, NOT ANNOTATED_CDS; Genomic_DNA.
PIR; A00704; JDVLVR.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000201; DNAPol_viral_N.
DR InterPro; IPR000477; RVse.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; RVT_1; 1.
DR ProDom; PD000814; DNAPol_viral_C; 1.
DR ProSITE; PS00878; RT_POL_1.
DR DNA replication; DNA-binding; DNA-directed DNA polymerase;
KW Endonuclease; Hydrolase; Multifunctional enzyme; Nuclease;
KW Nucleotidyltransferase; RNA-directed DNA polymerase; Transferrase.
FT DOMAIN 357 600 Reverse transcriptase.
SQ SEQUENCE 843 AA; 94400 MW; A6B2D490839C4E8B CRC64;
Alignment Scores:
Pred. No.: 1.18e-200 Length: 843
Score: 2807.00 Matches: 528
Percent Similarity: 98.33% Conservative: 3
Best Local Similarity: 97.78% Mismatches: 8
Query Match: 47.52% Indels: 2
DB: 1 Gaps: 0
US-10-761-006a-1 (1-3215) x DPOL_HPBVR (1-843)

QY 4 CACAAATCCACCAAGCTCTGTAGATCCAGGCTGAGGGGCTATATTTCTCTGCTGG 63
 Db |||||
 305 HisAsnIleProSerAlaAaGserGlnSerGluGlyProIlePheSerCysTrp 324
 QY TGCTCCAGTTCGGGAACAGTAAACCCCTGTTCGGACTACTCGCTCTCCCATATCGTCAAT 123
 Db |||||
 325 TrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsn 344
 QY 124 CTTCCTGAGACTGGGGACCCCTGACCGAATCATGAGAACACAACTCAGGATTCCTAGG 183
 Db |||||
 345 LeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArg 364
 QY 184 ACCCTGCTCGTGTACAGCGGGGTTTTCTCGTTGACAAAGATCCTCACATACCGCA 243
 Db |||||
 365 ThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThr 384
 QY 244 GAGTCTAGACTCG-GTGGACTTCTCTCAATTTCTAGGGGGAGCACCCACGCTGTTCCTG 302
 Db |||||
 385 GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-SerTr 404
 QY 303 GCCAAATTCGAGTCCCGAACCTCAATCACTCACCAACCTCTGTCTCTCAATTTGTC 362
 Db |||||
 404 pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe 424
 QY 363 CTGGCTATCGCTGATGTGTCTGGCGGTTTTATCATATTTCTCTTCATCCTGCTGCTAT 422
 Db |||||
 424 rTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMe 444
 QY 423 GCCTCATCTTCTTGTGGTTCCTCTGGACTTACCAAGGTATGTTGCCCGTTTGTCTCTAC 482
 Db |||||
 444 tProHisLeuLeuValValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerTh 464
 QY 483 TTCAGGAACATCAACACAGCAGCAGCGGCCATGCAAGACTGCAAGCTCTGTCTGCTCAAG 542
 Db |||||
 464 rSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSerAr 484
 QY 543 GAAACTCTAGTTTCTCTCTGT 602
 Db |||||
 484 gAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeuTyr 504
 QY 603 TTCCCATCCCATCATCTCTGGGCTTTTCGCAAGATTCCTATGGGAGTGGGCTCAGTCCGTT 662
 Db |||||
 504 rSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh 524
 QY 663 TCTCTCGCTCAGTTACTAGTCCCATTTGTTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 722
 Db |||||
 524 eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHisCys 544
 QY 723 TTTGGCTTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 782
 Db |||||
 544 sLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGly 564
 QY 783 GTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACCCCTAA 842
 Db |||||
 564 uSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAs 584
 QY 843 TAAACCAACGGTGGGCTACTCTCTTAATTCATGGGATATGTAATTTGGAAGTTGGGG 902
 Db |||||
 584 nLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrpGly 604
 QY 903 TACTTTACCGCAGACATATGTTACTTAAACTCAAGCAATGTTTTCGAAACTGCTCTCT 962
 Db |||||
 604 yThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuProVa 624
 QY 963 AAATAGACCTATTGATTGAAAGTATGTCAAGAAATTTGGGGTCTTTTGGGCTTTTCTGTC 1022
 Db |||||
 624 lAsnSerProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaI 644
 QY 1023 CCTTTTACCAATGTGGCTATCCTCGCTGTATGCTTTATATGATGATATCAATCTAA 1082
 Db |||||
 644 aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLys 664
 QY 1083 GCAGGCTTTTCTCTGCCCACTTACAAGGCTTTCTGTGTGTAAACAATAATCTGAACCT 1142

Db ||||| sGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsnLe 684
 QY 1143 TTATCCCGTTCGCCGCAACGGTCCGGTCTCTGCCAAGTGTTTTGTGTCAGCAACCCCCAC 1202
 Db ||||| uTyrProValAlaAaGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProTh 704
 QY 1203 TGGATGGGGCTTGGCCATAGGCATAGGCATAGGCATAGGCATAGGCATAGGCATAGGCATAGGC 1262
 Db ||||| rGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuPr 724
 QY 1263 GATCCATCTATCGGAACTCCTAGCAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1322
 Db ||||| olleHisThrAlaGluLeuLeuAlaAlaCysPheAlaAaGserArgSerGlyAlaLysLe 744
 QY 1323 TATCGAACCAGCAACTCTGT 1382
 Db ||||| uIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeuLe 764
 QY 1383 AGGTGTGTGTGCAACTGATCCTGCGGGGAGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1442
 Db ||||| uGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLe 784
 QY 1443 GAATCCCGGGAGCAGCCCGTCTCGGGGCGGTTTGGGGCTCTACCGTCCCTTCTTCATCT 1502
 Db ||||| uAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProleuLeuLeuLe 804
 QY 1503 GCGGTTCCGGCCGACACACGGGGCGCACTCTCTTTACGGGGTCTCCCGGTATGTGCTTC 1562
 Db ||||| uProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSe 824
 QY 1563 TCATCTCGCGGACCGGTGTGCACTTGTGCTTACCTCTGACGTCGATCGATGAGACCAACG 1620
 Db ||||| rHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 843

RESULT 9
 Q97975_HPBVO PRELIMINARY; PRT; 540 AA.
 AC Q97975;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE ORF (Fragment).
 OS Hepatitis B virus.
 OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
 OC Orthohepadnavirus.
 OX NCBI_TaxID=10407;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=95294549; PubMed=7775946;
 RA Uchida T., Gotoh K., Shikata T.;
 RT "Complete nucleotide sequences and the characteristics of two
 RT hepatitis B virus mutants causing serologically negative acute or
 RT chronic hepatitis B.";
 RL J. Med. Virol. 45:247-252 (1995).
 CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
 DR EMBL; D16666; BAA04073.1; -; Genomic DNA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR InterPro; IPR001462; DNAPol_viral_C.
 DR InterPro; IPR000477; RVase.
 DR Pfam; PF00336; DNA pol viral_C; 1.
 DR Pfam; PF00078; RVF_1; 1.
 DR Prodom; PD000814; DNAPol_viral_C; 1.
 DR PROSITE; PS50878; RT_POL; 1.
 FT NON_TER
 SQ SEQUENCE 540 AA; 60346 MW; 7BBE285359DD39C5 CRC64;

Alignment Scores:

Pred. No.:	1.2e-200	Length:	540
Score:	2806.00	Matches:	526
Percent Similarity:	98.34%	Conservative:	6
Best Local Similarity:	97.23%	Mismatches:	8
Query Match:	47.50%	Indels:	2
DB:	2	Gaps:	0

US-10-761-006A-1 (1-3215) x Q97975_HPBVO (1-540)

Qy	1	CTCCACAACTTCCACCAAGCTCTGCTAGATCCAGGGTGGGGGCTATATTTCTCTGC	60
Db	1	LeuHisAsnIleProProSerSerAlaAArgSerGlnSerGluGlyProIlePheSerCys	20
Qy	61	TGTTGGCTCAGTCCGACAGTAACCTGTTCCGACTACTCCCTCCCATATCGTC	120
Db	21	TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal	40
Qy	121	AATCTTCTCAGGAGTGGGACCTCGACCGAACATCGGAGAACACACATCAGGATTCCT	180
Db	41	AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleAArgIlePro	60
Qy	181	AGGACCCCTGCTGTTTACAGCGGGGTTTTCTGTTGACAAAGATCTCTCACATACC	240
Db	61	ArgThrProAlaAArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr	80
Qy	241	GCAGAGCTAGACTCTG- GTGGACTTCTCAATTTTCTAGGGGAGACCCAGTCTTC	299
Db	81	ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal- Se	100
Qy	300	CTGGCCAAATTCAGTCCCACTCCCAATCACTACCAACCTCTGTCCTCCAATTT	359
Db	100	rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe	120
Qy	360	GTCCTGGCTATCGTGGATGTCTGGGGCGTTTTATCATATTCCTCTCTATCTCTGTC	419
Db	120	uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl	140
Qy	420	TATGCTCATCTCTGTTGTTCTCTGGAATCAACAGGATGTTGGCGTTTGTCTCTC	479
Db	140	aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe	160
Qy	480	TACTTCCAGGAACATCAACACAGCAGCGGGCCATGCAAGACTCGACGACTCTGCTC	539
Db	160	rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe	180
Qy	540	AAGGAACTCTACGTTTCCCTCTGTTGCTGTACAAAACCTTCGAGCGGAACTGCATT	599
Db	180	rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe	200
Qy	600	GTATTCCCATCCATCATCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC	659
Db	200	uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr	220
Qy	660	GTTTCTCTGCTCAGTTTACTAGTCCATTTGTTTCTGAGTGTCTGAGGCTTTCCGCCA	719
Db	220	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi	240
Qy	720	CTGTTTGGCTTTCAGTTATATGATGATGTGTATTTGGGGCGAGTCTGTACACATCT	779
Db	240	sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe	260
Qy	780	TGAGTCCCTTTTACCTCTATTACAAATTTCTTTTCTTTTGGGTATACATTTAAACCC	839
Db	260	uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr	280
Qy	840	TAATTAACCAACCTTGGGCTACTCCCTTAACCTTCATGGATATGTAATTTGGAAGTTG	899
Db	280	oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr	300
Qy	900	GGGTACTTTTACCGCAGGAACATATTGTACTAAACCTCAAGCAATGTTTTTCGAAACTGCC	959
Db	300	pGlyThrLeuProGlnGluHisIleValLeuLeuLysLeuLysGlnCysPheArgLysLeuPr	320

Qy	960	TGTAATAGACCTATTGATTGGAAAGATATGTCAAAGGCTTTCTGTGTAACAAATATCTG	1019
Db	320	oValaAsnAArgProIleAspTrpLysValCysGlnAArgIleValGlyLeuLeuGlyPheAl	340
Qy	1020	TGCCCCCTTTTACAAATGTGGCTATCTGCTGCTGATGCTCTTATATGATGATATCAATC	1079
Db	340	aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe	360
Qy	1080	TAAGCAGGCTTTTCTCCTGCAACTTCAAGGCTTTCTGCTGTGTAACAAATATCTGAA	1139
Db	360	rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysGlnGlnTyrLeuHi	380
Qy	1140	CTTTTACCCCGTTCGCCGCAACGGTCCGCTCTCTGCCAAGTGTTCCTGACCGAACCCC	1199
Db	380	sLeuTyrProValAlaAArgGlnAArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr	400
Qy	1200	CATGATGGGCTTGGCCATAGCCATAGCCATAGCCATAGCCATAGCCATAGCCATAGCCAT	1259
Db	400	oThrGlyTrpGlyLeuAlaIleGlyHisAArgMetArgGlyThrPheValAlaProLe	420
Qy	1260	GCCGATCCATCTGCGGAACCTCTAGCAGCTTGTGTTGCTCGCAGCGCGTCTGGAGCAAA	1319
Db	420	uProIleHisThrAlaGluLeuLeuAlaCysPheAlaAArgSerArgSerGlyAlaLy	440
Qy	1320	ACTTATCGGAACCGACAACTCTGTGCTCTCTCGGAAATACACCTCTTCCATGGCT	1379
Db	440	sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe	460
Qy	1380	GCTAGGTGTGCTGCTGCACTGGATCTCTGCGGGAGCGTCTTGTCTACGTCCTCGGCG	1439
Db	460	uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl	480
Qy	1440	GCTGATCCCGCGGACGACCGCTCTGCGGGCGCTTTCGGGCTCTACCGTCCCTTCTTCA	1499
Db	480	aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi	500
Qy	1500	TCTGCCCTTCCGGCCGACACCGCGCGCTCTCTTTACGCGCTCTCCCGTATGTCGCC	1559
Db	500	sLeuProPheAArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr	520
Qy	1560	TTCATCTCTGCGGACCGTGTGCACTTCTGCTTACCTCTGCGACGCTCGCATGGAGCCACC	1619
Db	520	oSerHisLeuProSerArgValHisPheAlaSerProLeuHisValAlaTrpAArgProPr	540
Qy	1620	G 1620	
Db	540	o 540	

RESULT 10

Q5KR39_HPBVO PRELIMINARY; PRT; 843 AA.

ID Q5KR39; HPBVO PRELIMINARY; PRT; 843 AA.

AC Q5KR39; HPBVO PRELIMINARY; PRT; 843 AA.

DT 10-MAY-2005 (TrEMBLrel. 30, Created)

DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)

DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)

DE Polymerase.

GN Name=P;

OS Hepatitis B virus.

OC Viruses; Retro-transcribing viruses; Hepadnaviridae;

OC Orthohepadnavirus.

OX NCBI_TaxID=10407;

RN [1]

RN NUCLEOTIDE SEQUENCE.

RA Sun X., Rokounhara A., Tanaka E., Gad A., Mutou H., Matsumoto A.,

RA Yoshizawa K., Kiyosawa K.;

RT "Nucleotide Mutations Associated With Hepatitis B e Antigen

RT Negativity.";

RL J. Med. Virol. 76:170-175(2005).

RN [2]

RN NUCLEOTIDE SEQUENCE.

RA Rokuhara A.;

RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.

EMBL; AB198077; BAD86600.1; -; Genomic_DNA.

DR GO: 0003677; F: DNA binding; IEA.
 DR GO: 0003887; F: DNA-directed DNA polymerase activity; IEA.
 DR GO: 0004519; F: endonuclease activity; IEA.
 DR GO: 0016787; F: hydrolase activity; IEA.
 DR GO: 0004523; F: ribonuclease H activity; IEA.
 DR GO: 0003723; F: RNA binding; IEA.
 DR GO: 0003964; F: RNA-directed DNA polymerase activity; IEA.
 DR GO: 0016740; F: transferase activity; IEA.
 DR GO: 0006278; F: RNA-dependent DNA replication; IEA.
 DR InterPro: IPR001462; DNAPol_viral_C.
 DR InterPro: IPR00201; DNAPol_viral_N.
 DR InterPro: IPR000477; RVtse.
 DR Pfam: PF00336; DNA_pol_viral_C; 1.
 DR Pfam: PF00242; DNA_pol_viral_N; 1.
 DR Pfam: PF00078; RVt_1; 1.
 DR ProDom: PD000814; DNAPol_viral_C; 1.
 SQ SEQUENCE 843 AA; 94446 MW; 80CD75F48E81AC2F CRC64;

Alignment Scores:

Pred. No.:	1.4e-200	Length:	843
Score:	2806.00	Matches:	526
Percent Similarity:	98.15%	Conservative:	5
Best Local Similarity:	97.23%	Mismatches:	9
Query Match:	47.50%	Indels:	2
DB:	2	Gaps:	0

US-10-761-006a-1 (1-3215) x Q5KR39_HPBVO (1-843)

QY	1	CTCCACAACTTCCACCAAGCTCTGCTAGATCCAGGGTGGGGGCTATATTTCTCTGC	60
DB	304	LeuHisAsnIleProProSerSerAlaAtrpGlnSerGluGlyProIleLeuSerCys	323
QY	61	TGTGGCTCCAGTCCGGAACTAAACCTGTTCGACTACTGCTCTCCCATATCGTC	120
DB	324	TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal	343
QY	121	AATCTTCTGAGACTGGGACCTCGACCGACATCGGAGAACACACATCAGATTCT	180
DB	344	AsnLeuLeuGlnAspTrpGlyProCysThrGluHisGlyHisAsnIleArgIlePro	363
QY	181	AGGACCCCTGCTCGTGTATCAGCGGGGTTTCTCGTTGACAAAGATCTCAACAATACC	240
DB	364	ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr	383
QY	241	GCAGAGCTAGACTCTG-GTGGACTTCTCAATTTCTAGGGGAGACCCACGTGTC	299
DB	384	ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se	403
QY	300	CTGGCCAAATTCGACGTCGCCAACCCTCCATCACTCACCACCTCTGCTCCGAATT	359
DB	403	rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe	423
QY	360	GTCTCGCTATCGCTGATGTGTCGGGGTGTATCATATTCCTTCATCTCTGCTGC	419
DB	423	uSerTrpLeuSerLeuAspValSerAlaAlaPheThrHisIleProLeuHisProAlaAl	443
QY	420	TATGCTCATCTTCTTGTGGTTCTTCTGAGCTACCAAGTATGTCGCCGTTGTCCTC	479
DB	443	aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe	463
QY	480	TACTTCCAGGAACATCAACACAGACCGGGGCATGCAAGACTCGACGACTCTGCTC	539
DB	463	rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe	483
QY	540	AAGGAACTCTAGCTTCTCTGTTGCTGTGTACAAAACCTTCGGACGGAAACCTGACATT	599
DB	483	rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe	503
QY	600	GTATTCCTCCATCATCTCGGCTTTTCGAAAGATTCCTATGAGGATGGGCTCAGTCC	659
DB	503	uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr	523
QY	660	GTTTCTCTGGCTCAGTTTACTAGTCCATTGTTTCAGTGGTTCGTAGGGCTTCCCCCA	719

RESULT 11

Q7TDR3_HPBVO

ID Q7TDR3_HPBVO PRELIMINARY; PRT; 843 AA.

AC Q7TDR3;

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

Db	523	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi	543
QY	720	CTGTTTGGCTTTTCAGTTATATGATGATGTTGTTGGGGCGGAAGTCTGTCAACATCT	779
Db	543	sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe	563
QY	780	TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC	839
Db	563	uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr	583
QY	840	TAATAAAACCAACGTTGGGGTACTCCCTTAACTTCTGCGGATATGTAATTCGAAGTTG	899
Db	583	oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr	603
QY	900	GGTACTTTTACCGCAGGAACATATTTACTATAAATCAAGCAATGTTTTCGAAAACCTGCC	959
Db	603	pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr	623
QY	960	TGTAAATAGACCTATTGATTGGAAAGTATGTCAAAGATTGTGGGTCTTTTGGGCTTGC	1019
Db	623	oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl	643
QY	1020	TGCCCTTTTACACAATGTTGGCTATCTCGCTTGTATGCTTATATGATGATATCAATC	1079
Db	643	aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe	663
QY	1080	TAAGCAGGCTTTTCACTTTCTCGCAACTTACAGGCTTTTCTGTGTAACAATATCTGAA	1139
Db	663	rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs	683
QY	1140	CCTTTACCCGTTGCCCCGAAACGGTCCGCTCTGCCAAGTGTGTTGCTGACACCAACCC	1199
Db	683	nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr	703
QY	1200	CCTGATGATGGGCTGCGCATAGGCCATCAGCGCATGGCTGGAACTTTCTGCTCTCTCT	1259
Db	703	oThrGlyTrpGlyLeuAlaValGlyHisArgArgMetArgGlyThrPheValAlaProLe	723
QY	1260	GCGATCCCATCTCGCGAACTCTTAGCAGCTGTTTGTCTCGCAGCGGTCTCGAGCAA	1319
Db	723	uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaL	743
QY	1320	ACTTATCGGAAACGCAACTCTGTTGCTCTCTCTCGAAATACACCTCTCTTCATGGCT	1379
Db	743	sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe	763
QY	1380	GCTAGGCTGCTGCGCAACTGGATCTCGCGGACGCTCTTGTCTAGTCCCGTCGGC	1439
Db	763	uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl	783
QY	1440	GCTGAATCCCGGACGACCCGCTCTCGGGCGCTTTTGGGGCTCTACCGTCCCCCTTCTCA	1499
Db	783	aleuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi	803
QY	1500	TCTGCGTTTCCGCGCCGACCAACCGGGCGCACTCTCTTTTACGGGTCTCCCGGTATGTC	1559
Db	803	sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr	823
QY	1560	TTCTCATCTGCGGACCGGTGTCACCTTCCTTACCTCTGCGCTGCGATGGAGACCCACC	1619
Db	823	oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr	843
QY	1620	G 1620	
Db	843	O 843	

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polymerase protein.
 OS Hepatitis B virus.
 OC Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.
 ON NCBI_TaxID=10407;
 RX (1)
 RP NUCLEOTIDE SEQUENCE
 RA Liu C.-J., Chen P.-J., Lai M.-Y., Kao J.-H., Chen D.-S.;
 RL Submitted (Oct-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY167096; AA041314.1; -; Genomic DNA.
 DR GO: GO:0003677; F: DNA binding; IEA.
 DR GO: GO:0003887; F: DNA-directed DNA polymerase activity; IEA.
 DR GO: GO:0004519; F: endonuclease activity; IEA.
 DR GO: GO:0016787; F: hydrolase activity; IEA.
 DR GO: GO:0004523; F: ribonuclease H activity; IEA.
 DR GO: GO:0003723; F: RNA binding; IEA.
 DR GO: GO:0003964; F: RNA-directed DNA polymerase activity; IEA.
 DR GO: GO:0016740; F: transferase activity; IEA.
 DR GO: GO:0006278; P: RNA-dependent DNA replication; IEA.
 DR InterPro; IPR001462; DNapol_viral_C.
 DR InterPro; IPR00201; DNapol_viral_N.
 DR InterPro; IPR00477; RVTse.
 DR Pfam; PF00336; DNA_pol_viral_C; 1.
 DR Pfam; PF00242; DNA_pol_viral_N; 1.
 DR Pfam; PF00078; RVT_1; 2.
 DR ProDom; PD000814; DNapol_viral_C; 1.
 SQ SEQUENCE 843 AA; 94350 MW; E033A0F3BD9963AB CRC64;

 Alignment Scores:
 Pred. No.: 1.98e-200 Length: 843
 Score: 2804.00 Matches: 526
 Percent Similarity: 98.15% Conservative: 5
 Best Local Similarity: 97.23% Mismatches: 9
 Query Match: 47.47% Indels: 2
 DB: 2 Gaps: 0

 US-10-761-006a-1 (1-3215) x Q7TDR3_HPBV0 (1-843)
 QY 1 CTCACAACTCCACCAAGCTCTGCTAGATCCACGGGTAGGGGCTATATTTCTCTGC 60
 DB 304 LeuGlnAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
 QY 61 TGGTGGCTCCAGTTCGGAACAGTAACCTGCTCCGACTACTGCTCTCCCATATGCTC 120
 DB 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
 QY 121 AATCTTCTCGAGGACTGGGGACCTCGACCCAAATGAGAACACAAATCAGGATTCCT 180
 DB 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
 QY 181 AGGACCCCTGCTCGGTTCACAGGGGGGTTTCTCGTTGACAAATCCTCACATACC 240
 DB 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
 QY 241 GCAGGTCTAGACTCTG-GTGGACTTCTCAATTTCTAGGGGAGCACCCAGCTGTC 299
 DB 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
 QY 300 CTGGCAAAATTCGAGTCCCAACCTCCAATCACTCACCAACCTCTTGCTCTCAATTT 359
 DB 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
 QY 360 GTCCTGGCTATCGCTGGATGTCTGCGGGTATTCATATTCCTCTTCATCTCTGCTGC 419
 DB 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
 QY 420 TATGCTCATCTCTCTGTTGGTCTCTCTGACTACCAAGGTATGTTGCCCTTTGCTCTC 479
 DB 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaAlaArgLeuSer 463
 QY 480 TACTTCCAGGAACATCAACACACGACGGGGCCATGCAAGACCTGACGACTCCTGCTC 539
 DB 463 rThrSerArgAsnIleAsnHisGlnHisIleGlyAlaMetGlnAspLeuHisAspSerCysSe 483

QY 540 AAGGAAACTCTAGCTTTCCCTCTTGTGTGTGTACAAAACTTCGGACGGAACCTGCACTT 599
 DB 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
 QY 600 GTATTCCTCCATCCCATCTCTGGCTTTCCAGATTCCTATGGAGTGGGCTCAGTCC 659
 DB 503 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
 QY 660 GTTTCCTCTGGCTCAGTTACTAGTCCATTTCTCAGTGGTTCGTAGGCTTCCCCCA 719
 DB 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProH 543
 QY 720 CTGTTGGCTTTTCTAGTATATGATGATGATGGGTATTGGGGCGGAAGTCTCTACAACATCT 779
 DB 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
 QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTCTTTTGGTATATACATTTAAACCC 839
 DB 563 oGluSerLeuTyrThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
 QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACCTCAAGCAATGTTTTCGAAAACCTGCC 899
 DB 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
 QY 900 GGGTACTTTTACCAGCAACATATTGTACTATAAACTCAAGCAATGTTTTCGAAAACCTGCC 959
 DB 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
 QY 960 TGTAATAGACCTATTGATTTGGAAGATATGTCAAAGAAATTTGGGCTTTTGGGCTTTGC 1019
 DB 623 oValAsn***ProIleAspTrpLysValCysGlnLysIleValGlyLeuLeuGlyPheAl 643
 QY 1020 TGCCCCCTTTTACACAAATGCTCTCTGCTGATGCTCTTATATCATGATATACAAATC 1079
 DB 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
 QY 1080 TAAGACGGCTTTCACTTTCTCGCAACTTACAAGGCCTTTCTGTGTAAACAATATCTGAA 1139
 DB 663 rLysGlnAlaPheThrPheSerProThrTyrLysThrPheLeuCysLysGlnTyrLeuAs 683
 QY 1140 CTTTATCCCGTTCGCCGCAACGGTCCGGTCTCTGCCAAGTGTGTGCTGACGCAACCCC 1199
 DB 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
 QY 1200 CACTGATGGGCTTGGCCATAGCCCATAGCCATAGCCATGGCTGGAACCTTCTGGCTCTCT 1259
 DB 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 723
 QY 1260 GCCGATCCATCTCGGCACTCTAGCAGCTTGTGTCTCGCAGCCGGTCTGGAGCAAA 1319
 DB 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaL 743
 QY 1320 ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGGAATATACACCTCTCTTTCATGCT 1379
 DB 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
 QY 1380 GCTAGGTGTCTGCCAAGTCTGCTGCGGGACCTCTCTTCTGCTACGTCCTGCTCGCCG 1439
 DB 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
 QY 1440 GCTCAATCCCGGACGACCCGCTCTCGGGGCGGTTTGGGGCTCTACCGTCCCTTCTTCA 1499
 DB 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
 QY 1500 TCTGCCGTTCGGCGGACCGACCGGGCGGCACTCTCTTTTACCGGCTCTCCCGTATGTGCC 1559
 DB 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
 QY 1560 TTCTCATCTGCCGACCGTGTGCTCTGCTTCCCTTCCCTCTGACGCTCGCATGGAGACCA 1619
 DB 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaAlaTrpArgProPr 843

QY 1380 GCTAGGGTGTGCTGCCAAGTGGATCTCGCGGGAGCGTCTTTGTCTAGTCCGTCGGC 1439
 Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
 QY 1440 GCTGAATCCCGCGGACGACCGTCTCGGGGCCGTTTGGGGCTCTACCGTCCCGTCTTCA 1499
 Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuPheArgProLeuLeuHl 803
 QY 1500 TCTGCCGTTCCGCGCACGAGCGGCGCACCTCTCTTTACCGGTTCTCCCGTATGTGC 1559
 Db 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
 QY 1560 TTCTCATCTCGCGGACGCTGCACTTCGCTTACCTCTGCACGTGCATGGAGACACC 1619
 Db 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
 QY 1620 G 1620
 Db 843 o 843

RESULT 14

ID Q81165_HPBVO PRELIMINARY; PRT; 838 AA.
 AC Q81165;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE DNA polymerase.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxID=10407;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=subtype adr;
 RX MEDLINE=93096607; PubMed=1461746;
 RA Mukaiide M.;
 RT "The complete nucleotide sequence of hepatitis B virus, subtype adr
 RT (SRADR) and phylogenetic analysis.";
 RL Nucleic Acids Res. 20:6105-6105(1992).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=subtype adr;
 RA Uchida T., Aye T.T., Shimojima M., Gotoh K., Shikata T.;
 RT "Full-length nucleotide sequence of a hepatitis B virus (HBV) mutant
 RT isolated from a patient with acute hepatitis who did not exhibit
 RT serological markers for HBV infection.";
 RL Int. Hepatol. Commun. 2:70-73(1994).
 DR EMBL; D16665; BAA04072.1; -; Genomic_DNA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0004519; F:endonuclease activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR InterPro; IPR001462; DNAPol_viral_C.
 DR InterPro; IPR000201; DNAPol_viral_N.
 DR InterPro; IPR000477; RVase.
 DR Pfam; PF00336; DNA_pol_viral_C; 1.
 DR Pfam; PF00242; DNA_pol_viral_N; 1.
 DR Pfam; PF00078; RVT_1; 2.
 DR ProDom; PD000814; DNAPol_viral_C; 1.
 SQ SEQUENCE 838 AA; 93902 MW; A686F42940C3B806 CRC64;

Alignment Scores:

Pred. No.: 2,78e-200 Length: 838
 Score: 2802.00 Matches: 525
 Percent Similarity: 98.33% Conservative: 6
 Best Local Similarity: 97.22% Mismatches: 8
 Query Match: 47.44% Indels: 2
 DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x Q81165_HPBVO (1-838)
 QY 4 CACAACATTCCACCAAGCTCTGTAGATCCAGGGTGGGGCGCTATATTTCTCTGCTGG 63
 Db 300 HisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCysTrp 319
 QY 64 TGGCTCCAGTTCCGGAACAGTAAACCTGTTCCGACTACTGCTCTCTCCATATCGTCAAT 123
 Db 320 TrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsn 339
 QY 124 CTTCTCGAGAGCTGGGGACCCCTGCACCGAGATGGAGAACACAAACATAGGATTCCTAGG 163
 Db 340 LeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArg 359
 QY 184 ACCCTGCTCGTGTATACAGCGGGGTTTTCTCGTTGACAAAGAAATCCTCACATACCGCA 243
 Db 360 ThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThrThr 379
 QY 244 GAGCTAGACTCTG- GTGACTTCTCTCAATTTTCTAGGGGAGACCCACGTTGCTCTG 302
 Db 380 GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-SerTr 399
 QY 303 GCCAAATTCGCAGTCCCAACCTCCAATCACTCAACCACTCTCTGCTCTCAATTTGTC 362
 Db 399 pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeu 419
 QY 363 CTGGCTATCGCTGGATGTGTCTGCGCGGTTTTATCATATTCCTCTTCATCCTGCTCTAT 422
 Db 419 rTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAla 439
 QY 423 GCCTCATCTTCTGTGGTCTTCTGGACTACCAAGTATGTTGCCCGTTGCTCTCTAC 482
 Db 439 tProHisLeuLeuValGlySer-SerGlyLeuProArgTyrValAlaArgLeuSerSerTh 459
 QY 483 TTCCAGGACATCAACCCACGACCGGGCCATGCAAGACCTGCACGACTCTCTGCTCAAG 542
 Db 459 rSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSerAr 479
 QY 543 GAAACTCTACGTTTCCCTCTTGTGTGTACAAAACCTTCGGACGGAAACTGCACATTGA 602
 Db 479 gAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeuTy 499
 QY 603 TTCCCATCCATCACTCGGCTTTTCGCAAGATTCCTATGGGAGTGGGGCTCAGTCCGTT 662
 Db 499 rSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh 519
 QY 663 TCTCTCGGCTCAGTTTACTAGTGCCATTTGTTTCAGTGGTTCGTAGGGCTTTTCCCCACATG 722
 Db 519 eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCy 539
 QY 723 TTTGGCTTTTCAGTTATATGATGTGTGTATTTGGGGCGAAGTCTGTACAACATCTTGA 782
 Db 539 sLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGl 559
 QY 783 GTCCCTTTTACTCTATTTACCAATTTTCTTTGCTTTTGGGTATACATTAAACCTTAA 842
 Db 559 uSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAs 579
 QY 843 TAAACCAACGTTGGGGTACTCCCTTAACCTTCATGGGATATGTAATTTGGAAGTTGGGG 902
 Db 579 nLysThrLysArgTrpGlyTyr-SerLeuAsnPheMetGlyTyrValIleGlySerTrpGl 599
 QY 903 TACTTTACCGCAGGAACATATTGTACTAAACCTCAAGCAATGTTTTCGAAAACTGCCTGT 962
 Db 599 yThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuProVa 619
 QY 963 AAATAGACCTATTGATGGAAAGTATGTCAAAGATTTGGGTCTTTTGGGCTTCTGCTGC 1022
 Db 619 lAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl 639
 QY 1023 CCCTTTTACCAATGTGGCTATCTCCGCTTGATGCTCTTATATGATGATGATATACAATCTAA 1082

Db 639 aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyAlaCysileGlnSerLy 659
Qy 1083 GCAGGCTTTACATTTCTCGCAACTTACAGGCTTCTCTGTAACAAATATCTGAACCT 1142
Db 659 sGlnAlaPheThrPheSerProThrTyrTyAlaPheLeuCysGlnGlnTyrLeuHisLe 679
Qy 1143 TTACCCCGTTGCCGGCAACGGTCCGGTCTCTGCAAGTGTCTGCTGACCAACCCCCAC 1202
Db 679 uTyProValAlaAargGlnAargSerGlyLeuCysGlnValPheAlaAaspAlaThrProTh 699
Qy 1203 TGGATGGGCTTGGCCATAGCCATAGCCATAGCCATAGCTGGAACTTTCTGGCTCTCTGCC 1262
Db 699 rGlyTyrGlyLeuAlaileGlyHisAargMetArgGlyThrPheValAlaProLeuPr 719
Qy 1263 GATCCATCTGCGGAATCTCTAGCAGCTCTTTGCTCGCAGCGGTCTCGAGCAAACT 1322
Db 719 oIleHisThrAlaGluLeuAlaAaCysPheAlaAargSerArgSerGlyAlaAysLe 739
Qy 1323 TATCGGAACCGAACACTCTGTGTCTCTCTCGGAATACACTCTCTTCATGGCTGCT 1382
Db 739 uIleGlyThrAspAsnSerValValLeuSerArgLyTyThrSerPheProTyrProLeuLe 759
Qy 1383 AGGTGTGCTGCCAACTGGATCTCTGCGCGGACGTCTTTGCTGACCTCCGCTCGGCGCT 1442
Db 759 uGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyxValProSerAlaLe 779
Qy 1443 GAATCCCGGACGACCGCTCTCGGGCGCTTTGGGGCTCTACCGTCCCTCTTCTCATCT 1502
Db 779 uAsnProAlaAaspAspProSerArgGlyArgLeuGlyLeuTyArgProLeuLeuHisLe 799
Qy 1503 GCGGTTCCGGCGACCGGCGGCGACCTCTCTTTACGCGGTCTCCCGGTATGCGCTTC 1562
Db 799 uProPheArgProThrThrGlyArgThrSerLeuTyxAlaValSerProSerValProSe 819
Qy 1563 TCATCTCGCGGACCGTGTGCACTTCTGCTTCACTCTGACGCTGCGATGGAGACCCG 1620
Db 819 rHisLeuProSerArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 838

RESULT 15
O42041 HPBV0
ID O42041 HPBV0 PRELIMINARY; PRT; 843 AA.
AC O42041;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE DNA polymerase.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97160188; PubMed=9007704; DOI=10.1016/S0168-8278(96)80280-7;
RA Asahina Y., Enomoto N., Kurosaki M., Sakuma I., Izumi N., Marumo F.,
RA Sato C.;
RT "Sequential changes in full-length genomes of hepatitis B virus
RT accompanying acute exacerbation of chronic hepatitis B.";
RL J. Hepatol. 25:787-794 (1996).
DR EMBL; D50518; BAA23442.1; -; Genomic DNA.
DR EMBL; D50517; BAA23435.1; -; Genomic DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006740; F:transferase activity; IEA.
DR GO; GO:0006278; F:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNaPol_viral_C.
DR InterPro; IPR000201; DNaPol_viral_N.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.

DR Pfam; PF00078; RVT_1; 2.
DR ProDom; PD000814; DNaPol_viral_C; 1.
SQ SEQUENCE 843 AA; 94462 MW; 1BCCF79639BB140D CRC64;
Alignment Scores:
Pred. No.: 2,79e-200 Length: 843
Score: 2802.00 Matches: 525
Percent Similarity: 98.52% Conservative: 8
Best Local Similarity: 97.04% Mismatches: 7
Query Match: 47.44% Indels: 2
DB: 2 Gaps: 0
US-10-761-006a-1 (1-3215) x O42041_HPBV0 (1-843)
Qy 1 CTCACAAATTCACCAAGCTCTGTAGATCCAGGCTGAGGGGCTATATTTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerSerAlaAargSerGlnSerGluGlyProIlePheSerCys 323
Qy 61 TGGTGGCTCCAGTTCGGGAACAGTAAACCTGTTCGAGCTACTGCTCTCCCATATGCTC 120
Db 324 TrpTrpLeuGlnPheAargAsnSerLysProCysSerAspTyrCysLeuThrHisLeVal 343
Qy 121 AATCTTCTCGAGACTGGGGACCTCGCCGACATGGAGAACACACATCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleAargIlePro 363
Qy 181 AGGACCCCTGCTCGTGTACAGCGCGGTTTTCTCTGACCAAGATCTCTCAATATACC 240
Db 364 ArgThrProAlaAargValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
Qy 241 GCAGAGCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGCACCCACGTGTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
Qy 300 CTGGCCAAATTCGAGTCCCAACCTCCGATCACTCACCAACCTTTGTGCTCCCAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
Qy 360 GTCCTGGCTATCGCTGATGTGTGCGGGTTTTATCATATTCCTCTCTCATCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
Qy 420 TATGCTCTCATCTTCTTGTGTGCTTCTCTGGACTACCAAGTATGTTGCCGTTTGTCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaAargLeuSerSe 463
Qy 480 TACTTCCAGGAACATCAACACAGCAGCGGGCCATGACAGCTGCACGACTCTCTGCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisGluSerCysSe 483
Qy 540 AAGGAACTCTAGCTTTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
Db 483 rArgAsnLeuTyxValSerLeuLeuLeuLeuTyxLysThrPheGlyArgLysLeuHisLe 503
Qy 600 GTATTCCTCCATCCATCATCTCGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTTC 659
Db 503 uTyxSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy 660 GTTTCCTCTGGCTCAGTTACTAGTGCATTTGTTTCACTGTTGTTGTTGTTGTTGTTGTT 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValAargAlaPheProHi 543
Qy 720 CTGTTTGGCTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db 543 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 563
Qy 780 TGAGTCCCTTTTACCTCTATTTACCAATTTTCTTTTGTGTTGGGTATACATTTAAACCC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
Qy 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACCTTCACTGATGGATATGTAATTTGAGTTG 899
Db 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603

QY 900 GGCTACTTTACCGCAGACATATCTACTAATACTCAAGCAATGTTTCGAAACTGCC 959
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
 QY 960 TGTAAATAGACCTATTGATTGGAAGTATGTCAAAGAAATTGTGGTCTTTTGGGCTTTGC 1019
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
 QY 1020 TGCCCTCTTTTACACAATGTGGCTATCTGCTTGTATGCTCTTTATATGATGATATACAATC 1079
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
 QY 1080 TAGCAGGCTTTCACCTTCTCGGCAACTTACAGGCTTTCGTGTAAACAATATCTGAA 1139
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysGlnGlnTyrLeuHi 683
 QY 1140 CCTTTACCCCGTTGCCGCAACGGTCCGGTCTCTGCAAGTGTTCGCTGACGCAACCCC 1199
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 683 sLeuTyrProValAlaArgGlnArgSerGlyValCysGlnValPheAlaAspAlaThrPr 703
 QY 1200 CACTGATGGGCTTGCCCATAGCCATCAGCCATGGCTGGAACCTTTCTGGCTCCTCT 1259
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 723
 QY 1260 GCCGATCCATCTGCGGAACCTCTAGCAGCTTGTCTGCTCGCAGCGGCTCTGGAGCAA 1319
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 723 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLy 743
 QY 1320 ACTTATCGGAACCGACAACCTCTGTCTCTCTCGGAATACACCTCTCTTCCATGGCT 1379
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
 QY 1380 GCTAGGCTGTGCTGCCAATGGATCTCGCGGGACGTCCTTTGTCTACGTCCTCGTCGGC 1439
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
 QY 1440 GCTGAATCCCGCGGACGACCGCTCTCGGGCCGTTTGGGGCTCTACCGTCCCGCTTCTCA 1499
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
 QY 1500 TCTCGCGTTCCGGCCGACACGGGGCGCACCTCTCTTTACGGGCTCTCCCGTATGTGCC 1559
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
 QY 1560 TTCTCATCTGCGGACCGTGTGCACTTCGCTTCCCTCTGCAAGTCGCGATGGAGACCACC 1619
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
 QY 1620 G 1620
 Db |
 843 o 843

Search completed: December 1, 2005, 23:59:04
 Job time : 772.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: December 1, 2005, 23:16:00 ; Search time 94.5 Seconds
(without alignments)
5625.446 Million cell updates/sec

Title: US-10-761-006A-1
Perfect score: 5907
Sequence: 1 CTCACACATTCACCAAG.....CCTCAGGCCACGACGTGAA 3215

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-Q=/cgn2_1/USPTO.spool_p/US10761006/runat_01122005_113953_26895/app_query.fasta_1.3399
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10761006_@CGN 1 140 @runat_01122005_113953_26895 -NCFU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2927	49.6	843	2	US-09-719-528A-2
2	2927	49.6	843	2	US-10-209-264-2
3	2807	47.5	843	2	US-08-591-502B-45
4	2807	47.5	843	2	US-08-591-502B-59
5	2807	47.2	843	2	US-08-591-502B-47
6	2773	46.9	845	2	US-08-591-502B-46
7	2759	46.7	843	2	US-08-591-502B-48
8	2759	46.7	845	1	US-08-416-950-11
9	2759	46.7	845	2	US-08-469-830-11
10	2741	46.4	845	2	US-08-591-502B-11
11	2734.5	46.3	842	2	US-08-591-502B-51
12	2729.5	46.2	842	2	US-08-591-502B-50

13	2652	44.9	832	2	US-08-591-502B-61	Sequence 61, Appl
14	2646	44.8	832	2	US-08-591-502B-62	Sequence 62, Appl
15	2645	44.8	843	2	US-08-591-502B-56	Sequence 56, Appl
16	2643	44.7	832	2	US-08-591-502B-60	Sequence 60, Appl
17	2643	44.7	843	2	US-08-591-502B-55	Sequence 55, Appl
18	2637	44.6	832	2	US-08-591-502B-64	Sequence 64, Appl
19	2633	44.6	843	2	US-08-591-502B-53	Sequence 53, Appl
20	2615	44.3	843	2	US-08-591-502B-54	Sequence 54, Appl
21	2608	44.2	845	2	US-08-591-502B-57	Sequence 57, Appl
22	2607	44.1	832	2	US-08-591-502B-63	Sequence 63, Appl
23	2607	44.1	845	2	US-08-591-502B-58	Sequence 58, Appl
24	2607	44.1	845	6	5196194-17	Patent No. 5196194
25	2179	36.9	730	2	US-08-591-502B-49	Sequence 49, Appl
26	1496	25.3	400	2	US-09-719-528A-3	Sequence 3, Appl
27	1496	25.3	400	2	US-10-209-264-3	Sequence 3, Appl
28	1427	24.2	277	6	5164485-2	Patent No. 5164485
29	1423	24.1	281	2	US-09-247-890-10	Sequence 10, Appl
30	1423	24.1	281	2	US-09-724-969-10	Sequence 10, Appl
31	1423	24.1	281	2	US-09-724-852-10	Sequence 10, Appl
32	1365.5	23.1	481	2	US-08-591-502B-52	Sequence 52, Appl
33	1359	23.0	281	1	US-08-105-483-214	Sequence 214, App
34	1359	23.0	281	1	US-08-709-209-214	Sequence 214, App
35	1359	23.0	281	1	US-08-458-101-214	Sequence 214, App
36	1359	23.0	389	1	US-08-105-483-216	Sequence 216, App
37	1359	23.0	389	1	US-08-105-483-219	Sequence 219, App
38	1359	23.0	389	1	US-08-709-209-216	Sequence 216, App
39	1359	23.0	389	1	US-08-709-209-219	Sequence 219, App
40	1359	23.0	389	1	US-08-458-101-219	Sequence 219, App
41	1359	23.0	389	1	US-08-458-101-216	Sequence 216, App
42	1354	22.9	281	2	US-09-247-890-12	Sequence 12, Appl
43	1354	22.9	281	2	US-09-724-969-12	Sequence 12, Appl
44	1354	22.9	281	2	US-09-724-852-12	Sequence 12, Appl
45	1354	22.9	281	2	US-09-721-480-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-719-528A-2
; Sequence 2, Application US/09719528A
; Patent No. 6558675
; GENERAL INFORMATION:
; APPLICANT: Oon, Chong Jin
; Lim, Gek Keow
; Zhao, Yi
; Chen, Wei Ning

TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
USES THEREOF

NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ladas & Parry
STREET: 26 West 61 Street
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10023

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/719,528A
FILING DATE: 30-Apr-2001
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/SG98/00046
FILING DATE: 19-JAN-1998

ATTORNEY/AGENT INFORMATION:
NAME: Mass, Clifford J.
REGISTRATION NUMBER: 30,086
REFERENCE/DOCKET NUMBER: U-013109-7
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 708-1800

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 843 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-719-528A-2

Alignment Scores:

Pred. No.: 5,09e-259 Length: 843

Score: 2927.00 Matches: 540

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 49.55% Indels: 0

DB: 2 Gaps: 0

US-110-761-006A-1 (1-3215) x US-09-719-528A-2 (1-843)

```

QY 1 CTCACAACTTCACCAAGCTCTGTAGATCCAGGGTGGGGCTATATTTCTCTGC 60
Db
QY 304 LeuHisAsnIleProProSerSerAlaArgSerGlnGlyGluGlyProIlePheSerCys 323
Db
QY 61 TGTGGCTCCAGTTCGGGAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTC 120
Db
QY 324 TrpIrpLeuGlnPheArgAsnSerIleProCysSerAspIleVal 343
Db
QY 121 AATCTTCTCAGGACTGGGACCTGTCACCGAATCGGAGAACACCAATCAGGATTCCT 180
Db
QY 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGACCTCTCTCGTGTACAGCGGGGTTTTCTCGTTGACAAAGATCTCAATACC 240
Db
QY 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspIleValAsnThr 383
QY 241 GCAGCTCTAGACTCTGTGGACTCTCTCAATTTCTAGGGGGAGACCCAGCTGTTCC 300
Db
QY 384 AlaGluSerArgLeuIrpTrpThrSerLeuAsnPheLeuGlyGlyAlaProThrCysSer 403
QY 301 TGGCCAAAATTCGAGTCCCAACCTCCAACTACTCACCACCTCTTGTCTCTCAATTTG 360
Db
QY 404 TrpProIlePheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeu 423
QY 361 TCTCGCTATCGTGTGATGTCTGCGGCTTTTATCATATTCCTCTTCATCCTGCTGCT 420
Db
QY 424 SerTrpLeuSerLeuAspValSerAlaAlaPheThrHisIleProLeuHisProAlaAla 443
QY 421 ATGCTCTATCTCTGTTGTTCTTCTGGACTACCAAGTATGTTGCGGCTTGTCTCT 480
Db
QY 444 MetProHisLeuLeuValGlySerSerGlyLeuProArgIleValAlaArgLeuSerSer 463
QY 481 ACTTCCAGGAACATCAACACCAAGCAGCGGGCCCATGCAAGACCTGCACGACTCTCTCA 540
Db
QY 464 ThrSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSer 483
QY 541 AGAAACTCTACCTTCCCTCTGTTGCTGTACAAACCTTCGACGGAACCTGCACTTG 600
Db
QY 484 ArgLysLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeu 503
QY 601 TATTCCTCCATCATCTCTGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCCG 560
Db
QY 504 TyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPro 523
QY 661 TTTCTCTGCTCAGTTTACTAGTGCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 720
Db
QY 524 PheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHis 543
QY 721 TGTTCGCTTTCAGTTATATGATGATGTGTTATTCGGGGCGAAGTCTGTACAACTTT 780
Db
QY 544 CysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeu 563
QY 781 GAGTCCCTTTTACCTTATTACCAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTAAACCT 840

```

```

Db 564 GluSerLeuPheThrSerIleThrAsnPheLeuSerLeuGlyIleHisLeuAsnPro 583
QY 841 AATAAAACCAAAAGTGGGGCTACTCCCTTAACTTCATGGGATATGTAATGGAACTGG 900
Db 584 AsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrp 603
QY 901 GGTACTTTACCGGAGGAACATATTGTACTAAACCTCAAGCAATGTTTTCGAAACCTGCCT 960
Db 604 GlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPro 623
QY 961 GTAAATAGACCTATTGATTGGAAAGTATCTCAAGAAATGTTGGTCTTTTGGCTTTGCT 1020
Db 624 ValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAla 643
QY 1021 GCCCCTTTTACACAATGTGGCTATCCTCGCTTGTATGCTCTTATATGTCATGTATACAACT 1080
Db 644 AlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSer 663
QY 1081 AGCAGGCTTTCACCTTCTCGGCAACTTACAGGCTTCTGTGTAAACAATATCTGAAC 1140
Db 664 LysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsn 683
QY 1141 CTTTACCCCGTTGCGCGCAACGGTCCGCTCTCTGCCAAGTGTTCGCTCAGCAACCCCC 1200
Db 684 LeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPro 703
QY 1201 ACTGATGGGGCTTGCCCATAGGCCATCAGCCATGGCTGGAACCTTCTGGCTCCTCTG 1260
Db 704 ThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetAlaGlyThrPheLeuAlaProLeu 723
QY 1261 CGATCCATACCTCGGAACTCTAGCAGCTTGTGTTGCTCGCAGCGGCTCGAGCAAAA 1320
Db 724 ProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 743
QY 1321 CTTATCGGAACCAACTCTGTGCTCTCGGAAATACACCTCTCCCTTCCATGGCTG 1380
Db 744 LeuIleGlyThrAspAsnSerValValLeuSerArgLysThrSerPheProIleProLeu 763
QY 1381 CTAGGCTGTGCTGCCAACTGGATCTCGCGGGAGCTCTTGTCTACCTCCCGTCGGCG 1440
Db 764 LeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAla 783
QY 1441 CTGAATCCCGGAGACACCCGCTCTCGGGCCGTTTGGGGCTCTACGCTCCCTTCTCAT 1500
Db 784 LeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuHis 803
QY 1501 CTGCGCTTCCCGCCGACCAACCGCGGCGCACTCTCTTACCGCGCTCTCCCGTATGTCCT 1560
Db 804 LeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProTyrValPro 823
QY 1561 TCTCATCTCGCGGACCGTGTGCATCTTCCTTCCCTCTGACGCTGCGCATGAGCACCG 1620
Db 824 SerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 843

```

RESULT 2

US-10-209-264-2

; Sequence 2, Application US/10209264

; Patent No. 6787142

; GENERAL INFORMATION:

; APPLICANT: Oon, Chong Jin

; Lim, Gek Keow

; Zhao, Yi

; Chen, Wei Ning

; TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND

; USSES THEREOF

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Ladas & Parry

; STREET: 26 West 61 Street

; CITY: New York

; STATE: New York

; COUNTRY: USA

ZIP: 10023
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/209,264
FILING DATE: 31-Jul-2002
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/SG98/00046
FILING DATE: 19-JAN-1998
ATTORNEY/AGENT INFORMATION:
NAME: Mass, Clifford J.
REGISTRATION NUMBER: 30,086
REFERENCE/DOCKET NUMBER: U-013109-7
TELEPHONE: (212) 708-1800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2:

541 AGGAAACTCTACGTTTCCCTCTTGTTGCTGTACAAACCTTCGACGGAACATGCACTTG 600
|||||
484 ArgLysLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeu 503
601 TATTCACATCCCATCATCTGGGCTTTCGGAAGATTCCTATGGAGTGGCCCTCAGTCCG 660
|||||
504 TyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPro 523
661 TTTCTCTGCTCAGTTTACTAGTGCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCAC 720
|||||
524 PheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHis 543
721 TGTTCGCTTTTTCAGTTTATATGATGTGATGTTGGGGCGGAGTCTGTACAACTCTT 780
|||||
544 CysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLeu 563
781 GAGTCCCTTTTACCTCTATTACCAATTTTCTTTTCTTTGGGTATACATTTAAACCTT 840
|||||
564 GluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPro 583
841 AATAAACCACCAACGTTGGGCTACTCCCTTAACTTCATGGCATATGTAATTGGAAGTTGG 900
|||||
584 AsnLysThrLysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTyr 603
901 GGTACTTTTACCGCAGGAAACATATTGTACTAAAACTCAAGCAATGTTTTCGAAACCTCCT 960
|||||
604 GlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPro 623
961 GTAAATAGACTATTGATGCGAAAGTATGCAAGAAATGTTGGGTCTTTTGGGGCTTTGCT 1020
624 ValAsnArgProIleAspTyrLysValCysGlnArgIleValGlyLeuLeuGlyPheAla 643
1021 GCCCTTTTACACAACTGGCTATCCGCTTCATGCTTATATGATGATGATATCAATCT 1080
|||||
644 AlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSer 663
1081 AAGCAGGCTTTTCACTTTCTCGCAACTTCAAGGCTTTCTGTGTAAACAAATATCTGAAC 1140
664 LysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsn 683
1141 CTTTACCCCGTTCGCCGCAACGCTCCGCTCTGCGCAAGTGTGTCGACGCAACCCCT 1200
684 LeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPro 703
1201 ACTGGATGGGCTTGGCCATAGGCATCAGGCATGCTGCACTTCTGCTCTCTGCTCTCTG 1260
704 ThrGlyTyrGlyLeuAlaIleGlyHisGlnArgMetAlaGlyThrPheLeuAlaProLeu 723
1261 CCGATCCATCTCGCGAACTCTAGCAGCTTGTGTTGCTCGCAGCCGCTCTGGAGCAAAA 1320
724 ProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 743
1321 CTTATCGGAACCGCAACTCTGTTGCTCTCTCGGAAATACACTCTCTTCCATGGCTG 1380
744 LeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTyrLeu 763
1381 CTAGGTTGCTGCCAATCGATCTCGCGGGAGCTCTTGTCTAGTCCCGCTCTTCTTCAT 1440
764 LeuGlyCysAlaAlaAsnTyrPheLeuArgGlyThrSerPheValTyrValProSerAla 783
1441 CTGAATCCCGGGAGCAGCCGCTCTCGGGCGCTTGTGGGCTCTACCGCTCCCTCTTCTCAT 1500
784 LeuAsnProAlaAspAspProSerArgLysLeuGlyLeuTyrArgProLeuLeuHis 803
1501 CTGCGCTTTCGCGCCGACACCGGGCGCACCTCTCTTTACGGGCTCTCCCGCTATGCTGCT 1560
804 LeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProTyrValPro 823
1561 TCTCATCTGCGCGGACCGTGTGCTGCTTCCCTTCACTCTGACGCTGCGATGAGACCCG 1620
824 SerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTyrArgProPro 843

121 AATCTTCTCGAGGACTCGGGACCTCGCACCGCAACATGAGAACACAAATCAGGATTCCT 180
344 AsnLeuLeuGluAspTyrGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
181 AGGACCCCTGCTCGTTTACAGGCGGGTTCCTGTTGACAGAAATCCTCACATACC 240
364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
241 GCAGAGCTAGACTCTGGTGGACTCTCTCAATTTCTAGGGGAGCACCCACGCTGTTCC 300
384 AlaGluSerArgLeuTyrThrSerLeuAsnPheLeuGlyGlyAlaProThrCysSer 403
301 TSGCCAAAATTTCGAGTCCCACTCCCAATCCTCACTCACCACTCTGTCCTCCCAATTTG 360
404 TrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeu 423
361 TCCTGGCTATCGCTGATGTGTCTCGGGGTTTTATCATATTCCTTCATCCTGCTGCT 420
424 SerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAla 443
421 ATGCTCATCTCTTGTGGTCTTCTGAGCTACCAAGTATGTTGCCGTTGCTCTCT 480
444 MetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSer 463
481 ACTTCCAGGAACATCAACACACGACGCGGGGCATGCAAGACTGTCAGCACTCTCTGCTCA 540
464 ThrSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSer 483

Alignment Scores:
Pred. No.: 5, 09e-259 Length: 843
Score: 2927.00 Matches: 540
Percent Similarity: 100.00% Conservations: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.55% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1 (1-3215) x US-10-209-264-2 (1-843)

RESULT 3
US-08-591-502B-45
; Sequence 45, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; ; Lymphocyte Responses to Hepatitis B Virus

NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591.502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 45:
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 45:

US-08-591-502B-45
Alignment Scores:
Pred. No.: 5e-248 Length: 843
Score: 2807.00 Matches: 528
Percent Similarity: 98.33% Conservative: 3
Best Local Similarity: 97.78% Mismatches: 8
Query Match: 47.52% Indels: 2
DB: 2 Gaps: 0
US-10-761-006A-1 (1-3215) x US-08-591-502B-45 (1-843)

Qy	4	CACAACTTCACCAAGCTCTGTAGATCCAGGCTGAGGGGCTATATTTCTCTGCTGG	63
Db	305	HisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCysTrp	324
Qy	64	TGGCTCCAGTTCCGGAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATATCGTCAAT	123
Db	325	TrpLeuGlnPheArgAsnSerLysProCysSerAspPyrCysLeuThrHisIleValAsn	344
Qy	124	CTTCTCGAGGACTGGGGACCTTCACCCGAACATGGAGAACACACATCAGGATTCCTAGG	183
Db	345	LeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArg	364

Qy	184	ACCCCTGCTCGTGTACAGCGGGGTTTTCTCGTTGCAAGAATCCTCACAAATCCGCA	243
Db	365	ThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThrThr	394
Qy	244	GAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGGAGCACCCACGCTGTTCTCGT	302
Db	385	GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-SerTr	404
Qy	303	GCACAAATTCGAGTCCCAACCTCCAAATCCTCACCACCACTCTGTCTCTCCAAATTTGTC	362
Db	404	pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe	424
Qy	363	CTGGCTATCGTGGATGCTCTCGCGGTTTTATCATATTCCTCTTCATCTCTCTCTAT	422
Db	424	rTrpLeuSerLeuAspValSerAlaAlaPheThrHisIleProLeuHisProAlaIle	444
Qy	423	GCCTCATCTTCTTGTGTTCTTCTGACTACCAAGTATGTTGCCCGTTGTCTCTCTAC	482
Db	444	tProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerTh	464
Qy	483	TTCCAGGAACATCAACCAACAGCAGCGGGCCATGCAAGACCTGCAGCAGCTCTCTGCTCAAG	542
Db	464	rSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSerAr	484
Qy	543	GAACCTCTACGTTTCCCTCTTGTGCTGTACAAAACCTTCGGACGGAACTGCACATTGTA	602
Db	484	gAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeuTy	504
Qy	603	TTCCCATCCCATCATCTCTGGGCTTTTCGCAAGATTCCTATGGGAGTGGGCTCAGTCCGTT	662
Db	504	rSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh	524
Qy	663	TTCTCTGGCTCAGTTTACTAGTGCACATTTGTCAGTGGTTCGTAGGGCTTCCCCCAGCTG	722
Db	524	eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHisCy	544
Qy	723	TTTGGCTTTTCAGTTATATGATGATGTGTATTGGGGCGAAGTCTGTACAACATCTTGA	782
Db	544	sIleAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGl	564
Qy	783	GTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACCTTAA	842
Db	564	uSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAs	584
Qy	843	TAAACCAAAAGTTGGGGCTACTCCCTTAACTTCATCGGATATGTAATTTGGAAGTTGGG	902
Db	584	nLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrpGl	604
Qy	903	TACTTTACCGCAGGAACATATTGTACTAAACCTCAAGCAATGTTTTCGAAAACTGCCTGT	962
Db	604	yThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuProVa	624
Qy	963	AAATAGACCTATTGATTGGAAAGTATGTCAAAGAAATTGTGGGTCTTTTGGGCTTTCTGC	1022
Db	624	lAsnSerProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl	644
Qy	1023	CCCTTTTACAAATATGGGCTATCTCGCTTGATGCTCTTATATGATGATATACAAATCTAA	1082
Db	644	aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLy	664
Qy	1083	GCAGGCTTTCACCTTCTCGCCCACTTACAGGCTTTCTGTGTAAACAATATCTGAACCT	1142
Db	664	sGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsnLe	684
Qy	1143	TTACCCCGTTTGGCCGCAACGGTCCCGTCTCTCGCAAGTGTGTTTGTCTGACGCAACCCAC	1202
Db	684	uTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProTh	704
Qy	1203	TGATGGGCTTGGCCATAGGCCATCAGCGCATGGGCTAGGAACTTTCTGGCTCTCTGCC	1262
Db	704	rGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLeuPr	724
Qy	1263	GATCCATCTGCGGAACCTCTACGACGCTTGTGTTTGTCTCGCAGCGGCTCTGGAGCAAACT	1322

Db 724 oileHisThAlaGluLeuLeuAlaAaCysPheAlaArgSerArgSerGlyAlaLysLe 744
Qy 1323 TATCGGAACGACAACTCTGTGTCTCTCTCGGAATATACACCTCTTCATGGCTGCT 1382
Db 744 uileGlyThrAspAsnSerValValLeuSerArgLysThrSerPheProThrLeuLe 764
Qy 1383 AGGTGTCTGCCAACCTGATCTCGCGGGAGCTCTTTGTCTACGTCCCGTCGGGCT 1442
Db 764 uGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTrpValProSerAlaLe 784
Qy 1443 GAATCCCGCGACACCGCTCTCGGGCGCTTTGGGCTCTACCGTCCCTTCTTCATCT 1502
Db 784 uAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTrpArgProLeuLeuLe 804
Qy 1503 GCCCTTCGGCGACACCGGGCGACCTCTCTTTAGCGGTCTCCCGGTATGGCTTC 1562
Db 804 uProPheArgProThrThrGlyArgThrSerLeuTrpAlaValSerProSerValProSe 824
Qy 1563 TCATCTCGCGACCGTGTGCACTTCGCTTCACTCTGCAGCTGCATGGAGACCG 1620
Db 824 rHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaItrpArgProPro 843

RESULT 4
US-08-591-502B-59
; Sequence 59, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESS: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-08-591-502B-59

Alignment Scores: 5e-248 Length: 843
Pred. No.: 2807.00 Matches: 527
Score: 98.34% Conservative: 5
Best Local Similarity: 97.41% Mismatches: 8
Query Match: 47.52% Indels: 2
DB: 2 Gaps: 0
US-10-761-006A-1 (1-3215) x US-08-591-502B-59 (1-843)
Qy 1 CTCACAAATTCCACCAAGCTCTGTAGATCCAGGGTGGGGCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
Qy 61 TGTGTGCTCCAGTTCCGGAACAGTAAACCTGTTCGACTACTGCTCTCTCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
Qy 121 AATCTTCTCGAGACTGGGGACCTCGCACCGAATGGAGAACACACATCAGGATTCTCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
Qy 181 AGGACCTCTCTGTGTACAGCGGGGTTTTCTCGTTGACAGATCTCTCAATATACC 240
Db 364 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 383
Qy 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGAGACCCACCTGTTTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
Qy 300 CTGCGCAAAATTCGAGTCCCACTCCCAATCACTCACCAACCTCTGTCTCTCAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
Qy 360 GTCTGTGCTATCGTGTGATGTGTCTGGGGGTTTTTATCATATTTCTCTTCTCATCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
Qy 420 TATGCTCTATCTTCTGTGTCTCTCTGACTACCAAGGTATGTTCCTGCTCTCTCTCTCT 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
Qy 480 TACTTCCAGGAACATCAACACGACGCGGGCCATCAAGACCTCGACGACTCTCTGCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 483
Qy 540 AAGGAAACTCTAGCTTTCCCTCTTGTGTGTGTATACAAACCTTCGACGGAACCTGCAC 599
Db 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
Qy 600 GTATTTCCATCCCATCATCTCTGGGCTTTCCGAAGATTCCTATGGAGTGGGCTCAGTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy 660 GTTTCTCTGGCTCAGTTTACTAGTGCATTTGTTCAGTGTGTCTGTAGGGCTTTTCCCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 543
Qy 720 CTGTTTGGCTTTTCAAGTTATATGGATATGTGGTATTTGGGGCGGAGTCTGTACAACATCT 779
Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
Qy 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
Qy 840 TAATAAAACCAAAAGTGGGGCTACTCCCTTTAACTTCATGGGATATGTAATGGAGTTG 899
Db 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
Qy 900 GGGTACTTTTACCAGGAGAACATATGTGTACTAAACTCAAGCAATGTTTTTCGAAACTGCC 959

603 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACCTATTGATTGGAAAGTATCTCAAAGAAATGTCGGTCTTTTGGGCTTTC 1019
Db 623 oValaenArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCTTTTACACAATGTGGCTATCTGCTGCTGTATGTCCTTTATATGCATGTATACAATC 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAGCAGGCTTTTCACTTTCTCGCAACTTACAAAGGCTTTTCTGTGTAACAATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 683
QY 1140 CCTTTACCCCGTTGCGCGGCAAGCTCGGTCTCTGCAAGTGTGTTGCTGACCAACCC 1199
Db 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
QY 1200 CACTGGATGGGCTTGCCCATAGCCATCAGCGCATGGCTGGAACTTTCTGGCTCTCT 1259
Db 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 723
QY 1260 GCGATCCATPACTGCGGAACCTCTAGCAGCTTGTGTTGCTCGCAGCGGCTCTGGAGCAA 1319
Db 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaL 743
QY 1320 ACTTATCGGAACGACACTCTGTTGCTCTCTCGGAATACACCTCTCTTCCATGGCT 1379
Db 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
QY 1380 GCTAGGTGTGTCGCAACTGGAATCTCGCGGAGCTCTCTTGTCTAGTCCCGTCGCGC 1439
Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
QY 1440 GCTGAATCCCGCGGACGACCCGCTCTCGGGCCGTTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
QY 1500 TCTGCGGTTCCGCGGACGACGCGGCGACCTCTCTTTTACGCGGTCTCCCGTATGTGC 1559
Db 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
QY 1560 TTTCTATCTGCGCGGCGGTGTGACCTTGTGCTTACCTCTGACGTGCGATGGAGACCC 1619
Db 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaAlaTrpArgProPr 843
QY 1620 G 1620
Db 843 O 843

RESULT 5

US-08-591-502B-47
; Sequence 47, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B

FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-08-591-502B-47
Alignment Scores:
Pred. No.: 2,75e-246 Length: 843
Score: 2788.00 Matches: 523
Percent Similarity: 97.9% Conservatives: 7
Best Local Similarity: 96.6% Mismatches: 10
Query Match: 47.2% Indels: 2
DB: 2 Gaps: 0
US-10-761-006A-1 (1-3215) x US-08-591-502B-47 (1-843)

QY 1 CTCACAACATTCACCAAGCTCTGCTAGATCCAGGGTGAGGGGCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
QY 61 TGTGTGCTCCAGTTCGGGACAGTAAACCTCTTCGACTACTGCTCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTCGAGACTGGGGACCTCGCACCAACATGGAGAACACAAACATCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTCGTGTATACAGCGGGGTTTTTCTCGTTGACAAGAATCCTCAATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGTCTAGACTCTG-GTGCACCTCTCTCAATTTCTAGGGGAGCACCCACCTGTTTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY 300 CTGGGCAAAATTCGCAGTCCCCAACCTCCAATCACTCAACCACTCTTCTCTCCAATTT 359
Db 403 TrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY 360 GTCTGGCTATCGCTGGATGCTCTGCGCGGTTTTTATCATATTCCTTTCATCTCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
QY 420 TATGCTCATCTTCTGTGTTGTTCTCTGAGTACCAAGGTATGTGCCGTTCTGCTTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACACACGACGCGGGGCCATGCAAGACCTGACAGCTCTCTGCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483

QY 540 AAGAACTCTAGCTTCCCTCTCTGTTGCTGTACAAACCTTCGGACGGAAACTGCACCTT 599
 Db |||||
 QY 483 rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
 Db |||||
 QY 600 GTATTCCTCCATCCCATCTCGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
 Db |||||
 QY 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
 Db |||||
 QY 660 GTTTCCTCTGCTCAGTTTACTAGTGCATTTGTTTCAGTGTGCTAGGCTTTCCCCCA 719
 Db |||||
 QY 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
 Db |||||
 QY 720 CTGTTGGCTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 779
 Db |||||
 QY 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
 Db |||||
 QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTCTGTTGGGTATACATTTAAACCC 839
 Db |||||
 QY 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
 Db |||||
 QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACTTCATGGATATATGTAATTTGAAGTTG 899
 Db |||||
 QY 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
 Db |||||
 QY 900 GGGTACTTTACCGAGGAACATATGTACTAACTCAAGCAATGTTTTCGAAACTGCC 959
 Db |||||
 QY 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
 Db |||||
 QY 960 TGTAAATAGACTATTGATGGAAGTATGTCAAGAAATTCGGTCTTTTGGGCTTTGC 1019
 Db |||||
 QY 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
 Db |||||
 QY 1020 TGCCCTTTTACAAATGTGCTATCTCTGCTTGCCTTGCCTTTATATGATGATATACATC 1079
 Db |||||
 QY 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
 Db |||||
 QY 1080 TAAGCAGCTTTTCACTTCTCGCAACTTACAGCCCTTCTGTGTAAACAATATCTGAA 1139
 Db |||||
 QY 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysGlnGlnTyrLeuHi 683
 Db |||||
 QY 1140 CTTTACCGCTGTCGGGCAACGCTCGGCTCTCTCGCAAGTGTTCGTGACGCAACCCC 1199
 Db |||||
 QY 683 sLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
 Db |||||
 QY 1200 CACTGTGATGGGCTTGGCCATAGGCCATCAGCGCATGCTGGAACCTTTCTGGCTCTCT 1259
 Db |||||
 QY 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValValProLe 723
 Db |||||
 QY 1260 GCGATCCATACCTGGGAACCTCTAGAGCTTGTTCCTCGCAGCCGCTCGGAGCAA 1319
 Db |||||
 QY 723 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaAspArgSerGlyAlaLy 743
 Db |||||
 QY 1320 ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGGAATACACCTCTCTTCCATGCT 1379
 Db |||||
 QY 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
 Db |||||
 QY 1380 GCTAGGCTGTGTCGCAACTGATCCTGCGGGGACGTCCTTTGTCTACGTCGCGTCCGC 1439
 Db |||||
 QY 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
 Db |||||
 QY 1440 GCTGAATCCCGGACGACCGCTCTCGGGCGGCTTGGGCTCTACCGTCCCTCTCTTCA 1499
 Db |||||
 QY 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuSe 803
 Db |||||
 QY 1500 TCTGCGCTTCCGGCGACACCGGCGGACCTCTCTTACGCGGCTCTCCCGTATGTCGC 1559
 Db |||||
 QY 803 rLeuProPheGlnProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
 Db |||||
 QY 1560 TTCTCATCTGCGGACCGCTGTGCTCTCTCTCTGCACTCTGCACTGCGATGAGACCCACC 1619
 Db |||||
 QY 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTatPrArgProPr 843
 Db |||||

QY 1620 G 1620
 Db 843 o 843
 RESULT 6
 US-08-591-502B-46
 ; Sequence 46, Application US/08591502B
 ; Patent No. 6607727
 ; GENERAL INFORMATION:
 ; APPLICANT: Chisari, Francis V.
 ; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
 ; Lymphocyte Responses to Hepatitis B Virus
 ; NUMBER OF SEQUENCES: 99
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/591,502B
 ; FILING DATE: 20-May-1996
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/749,540
 ; FILING DATE: 26-AUG-1991
 ; APPLICATION NUMBER: US 07/935,898
 ; FILING DATE: 26-AUG-1992
 ; APPLICATION NUMBER: US 08/100,870
 ; FILING DATE: 02-AUG-1993
 ; APPLICATION NUMBER: WO PCT/US94/08685
 ; FILING DATE: 01-AUG-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Weber, Ellen Lauver
 ; REGISTRATION NUMBER: 32,762
 ; REFERENCE/DOCKET NUMBER: 014740-000230US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 576-0200
 ; TELEFAX: (415) 576-0300
 ; INFORMATION FOR SEQ ID NO: 46:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 845 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: <Unknown>
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; FEATURE:
 ; NAME/KEY: Modified-site
 ; LOCATION: 16..18
 ; OTHER INFORMATION: /product= "OTHER"
 ; /note= "Xaa = unknown"
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 46:
 US-08-591-502B-46
 Alignment Scores:
 Pred. No.: 6,52e-245 Length: 845
 Score: 2773.00 Matches: 519
 Percent Similarity: 97.04% Conservative: 6
 Best Local Similarity: 95.93% Mismatches: 15
 Query Match: 46.94% Indels: 2
 DB: 2 Gaps: 0
 US-10-761-006A-1 (1-3215) x US-08-591-502B-46 (1-845)
 QY 1 CTCACAAACATTCCACCAAGCTCTGCTAGATCCCAAGGCTAGGGGCTATATTTCTTCGC 60
 Db 306 LeuHisAsnIleProSerSerAlaArgProGlnSerGluGlyProIleLeuSerCys 325

QY 61 TGTTGGCTCCAGTTCGGGAACAGATAACCTGTTCGACTACTGCTCTCCCATATCGTC 120
 Db 326 TrpTrpLeuGlnPheArgAenSerLysProCysSerAspTyrCysLeuThrHisIleVal 345
 QY 121 AATCTTCTGAGACTGGGACCTCGACCAATGAGAACACACATCAGGATTCCT 180
 Db 346 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyHisAsnIleArgIlePro 365
 QY 181 AGGACCCCTGCTCGTGTACAGCGGGGTTTTCTGTTGACAGAATCTCTCAATATACC 240
 Db 366 ArgThrProIleAlaValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 385
 QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGGAGCACCCACGTGTTCT 299
 Db 386 ThrGluSerThrLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 405
 QY 300 CTGGCCAAATTCGCAGTCCCAACCTCCAACTCACTCACCACCTCTGTCTCCCAATTT 359
 Db 405 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 425
 QY 360 GTCTGGCTATCGTGTGATGTCGTGGGGGTTTTATCATATTCCTTCTCATCTCTGCTGC 419
 Db 425 uSerTrpLeuSerLeuAspValSerAlaIaPheTyrHisIleProLeuHisProIleAla 445
 QY 420 TATGCCTCATCTTCTGTTGTTCTTCTGGACTACCAAGGTATGTTGCCCGTTTGTCTCTC 479
 Db 445 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValValCysLeuSerSe 465
 QY 480 TACTTCAGGAACATCMACACAGACGCGGGCCATGCAAGACTCGACGACTCTCTGCTC 539
 Db 465 rThrSerLysAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 485
 QY 540 AAGGAAACTCTAGCTTTCCTCTGTTGCTGTACAAACCTTCGGACGGAAACCTGCACATT 599
 Db 485 rArgAsnLeuTyrValSerLeuPheLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 505
 QY 600 GTATTCCTCCATCCATCTCTGGGCTTTCGCAAGATTCTCTATGGAGTGGGCTCAGTCC 659
 Db 505 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 525
 QY 660 GTTTCCTGGCTCAGTTTACTAGTGCATTTGTTGAGTGGTTCGATAGGCTTCCCCCA 719
 Db 525 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 545
 QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 779
 Db 545 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 565
 QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTCTTTGTTGTTGTTGTTGTTGTTGTTGTT 839
 Db 565 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 585
 QY 840 TAATAAACCAAGTTCGGGCTACTCCCTTAACCTTACCTGAGATGATGATTCGAAGTTG 899
 Db 585 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlyCysTr 605
 QY 900 GGGTACTTTTACCAGGAACATATTGACTAAACCTCAAGCAATGTTTTCGAAACTGCC 959
 Db 605 pGlyThrLeuProGlnGluHisIleValLeuLysIleValLeuGlnCysPheArgLysLeuPr 625
 QY 960 TGTAATAGACCTATTGATTTGGAAGTATGTCAAAGAAATTTGGGTCTTTTGGGCTTTGC 1019
 Db 625 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 645
 QY 1020 TGCCCTTTTACAAATGTCGCTATCTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1079
 Db 645 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 665
 QY 1080 TAAGACGCTTTCATCTTCTCGCAACTCAACAGGCTTTCTGTTGTAACAAATATCTGAA 1139
 Db 665 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuLeuCysLysGlnTyrLeuAs

QY 1140 CCTTATACCCCGTTCGGCGCAACGGTCCGGTCTCTGCGCAAGTGTGTTGCTGACGCAACCCC 1199
 Db 685 nLeuTyrProValAlaAraGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 705
 QY 1200 CACTGATCGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACTTTCTGGCTCCTCT 1259
 Db 705 oThrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 725
 QY 1260 GCCGATCCATCTAGCGAACTCTAGCAGCTTGTGTTGCTCGCAGCCGGTCTGGAGCAAA 1319
 Db 725 uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 745
 QY 1320 ACTTATCGAACCGCAACTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1379
 Db 745 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 765
 QY 1380 GTTAGGCTGCTGCGCAACTGATCTGCGCGGACGCTCTTTTACGGGCTCTCCCGTCCGCGC 1439
 Db 765 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 785
 QY 1440 GCTGAATCCCGCGGACGACCCGCTCTCGGGCGGCTTTGGGGCTCTACCGTCCCTTCTTCA 1499
 Db 785 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 805
 QY 1500 TCTGCCGTTCCGGCCGACCCAGCGGGCGGACCTCTCTTTACGGGCTCTCCCGTATGTGCC 1559
 Db 805 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 825
 QY 1560 TTCTCATCTGCGGACCGGTGCACCTTCCTTCACCTTCGACGCTGCGACGTGGAGACCAACC 1619
 Db 825 oSerHisLeuProAspArgValHisPheProSerProLeuHisValAlaTrpArgProPr 845
 QY 1620 G 1620
 Db 845 o 845
 RESULT 7
 US-08-591-502B-48
 ; Sequence 48, Application US/08591502B
 ; Patent No. 6607727
 ; GENERAL INFORMATION:
 ; APPLICANT: Chisari, Francis V.
 ; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
 ; Lymphocyte Responses to Hepatitis B Virus
 ; NUMBER OF SEQUENCES: 99
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/591,502B
 ; FILING DATE: 20-May-1996
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/749,540
 ; FILING DATE: 26-AUG-1991
 ; APPLICATION NUMBER: US 07/935,898
 ; FILING DATE: 26-AUG-1992
 ; APPLICATION NUMBER: US 08/100,870
 ; FILING DATE: 02-AUG-1993
 ; APPLICATION NUMBER: WO PCT/US94/08685
 ; FILING DATE: 01-AUG-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Weber, Ellen Lauver
 ; REGISTRATION NUMBER: 32,762

REFERENCE/DOCKET NUMBER: 014740-000230US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:

LENGTH: 843 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 48:

US-08-591-502B-48

Alignment Scores:
Pred. No.: 1.25e-243 Length: 843
Score: 2759.00 Matches: 519
Percent Similarity: 97.23% Conservative: 7
Best Local Similarity: 95.93% Mismatches: 14
Query Match: 46.71% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x US-08-591-502B-48 (1-843)

```

Qy 1 CTCCACAACTTCCACCAAGCTCTGCTAGATCCAGGCTGAGGGCCCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProSerCysAlaArgSerGlnSerGluGlyProIleSerSerCys 323
Qy 61 TGTGGGCTCCAGTTCGCGAAGCAAGTAAACCTGTTCCGACTACTCCCTCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerGluProCysSerAspTyrcysLeuThrHisIleVal 343
Qy 121 AATCTTCTCGAGAGCTGGGACCTCGACCCGAAATGATGAGAACACATCAGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
Qy 181 AGGACCTCTGCTGTTACAGGGGGGTTTCTGTTGACAAAGATCCCTCACAAATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspIleAsnProHisAsnThr 383
Qy 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCAATTTTCTAGGGGGAGCACCCAGCTGTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
Qy 300 CTGGCAAAATTCGAGTCCCAACCTCAATCACTACCAACCTTGTGCTCTCAATTT 359
Db 403 rTrpProIysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
Qy 360 GTCTGGCTATCGTGGATGCTCTGCGCGTCTTATCATATTCCTTCTCATCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
Qy 420 TATGCTCATCTCTTGTGTTCTTCTGGACTACCAAGGATGTTGCGCGTGTGCTCCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
Qy 480 TACTTCCAGGAACATCAACACACGACGCGGGCCATGCAAGCTGACGACTCTGCTGC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
Qy 540 AAGGAACTCTACGTTTCCCTCTGCTGCTGACAAACCTTCGGACGGAACCTGCACTT 599
Db 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgGlyLeuHisIle 503
Qy 600 GTATTCCCATCCATCATCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy 660 GTTCTCTCGGCTCAGTTTACTAGTCCCAATTTGTTTCAAGTGTCTAGGCTTTCCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 543
Qy 720 CTGTTTGGCTTTCAGTTATATGATGATGTGTATTTGGGGCGAAGTCTGTACACATCT 779

```

```

Db 543 sCysLeuAlaPheSerTyMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
Qy 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTCTGGGTATACATTTAAACCC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
Qy 840 TAATAAAACCAACAGTTGGGGCTACTCCCTTAACTTCAATTCATGGGATATGTAATGGAAGTTG 899
Db 583 oHisLeuThrLysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
Qy 900 GGCTACTTTACCGCAGGAACATATTGTATAAACTCAACCAATGTTTTCGAAAACCTGCC 959
Db 603 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 623
Qy 960 TGTAAATAGACCTATTGATTGGAAAGTATCTCAAGAAATGTGGGTCTTTTGGGCTTTGC 1019
Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Qy 1020 TGCCCTTTTACACAATGTGGCTATCTGCTCTGATGCTTATATCATGATGATATCAATC 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
Qy 1080 TAAGCAGCTTTTCACTTTCTGCCCAACTTACAAGGCTTTTCTGTGTAAACAAATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
Qy 1140 CTTTACCTCCCGTTCGCGCAACGGTCTCTGCAAGTGTCTGTCACGCAACCC 1199
Db 683 sLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
Qy 1200 CACTGGATGGGCTTGGCCATAGCCATAGCCATAGCCATAGCCATAGCCATAGCCATAGCCAT 1259
Db 703 oThrGlyTrpGlyLeuAlaIleGlyGlnSerGlyMetArgGlyThrPheValAlaProLe 723
Qy 1260 GCCGATCCATATCTGCGAAGTCTCTAGCAGCTTGTGTTGCTCGCAGCGGCTCTGAGCNA 1319
Db 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 743
Qy 1320 ACTTATCGGAACCGACAACTCTGTTGCTCTCTCGGAAATACACCTCTTTCATGCT 1379
Db 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpue 763
Qy 1380 GCTAGGCTGTGCTGCCAACTGGATCTCGCGGGAGCTCTTGTCTACGTCCTGCTCGC 1439
Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
Qy 1440 GCTGAATCCCGCGACGACCGCTCTCGGGGCGGCTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
Qy 1500 TCTGCGCTTCGCGCGACCGACCGGCGGCGACCTCTTTCACGCGGTCTCCCGTATGTC 1559
Db 803 sLeuProPheArgProThrThrGlyArgAlaSerLeuTyrAlaValSerProSerValPr 823
Qy 1560 TTCTCATCTGCGGACCGCTGTGCACTTCTGCTTCACTCTGACGCTGCGATGAGACCA 1619
Db 823 oSerHisLeuProValArgValHisPheAlaSerProLeuHisAlaIleTrpArgProPr 843
Qy 1620 G 1620
Db 843 o 843

```

RESULT 8

US-08-416-950-11

; Sequence 11, Application US/08416950

; Patent No. 5780036

; GENERAL INFORMATION:

; APPLICANT: CHISARI, Francis V.

; TITLE OF INVENTION: PEPTIDES FOR INDUCING CYTOTOXIC T

; TITLE OF INVENTION: LYMPHOCYTE RESPONSES TO HEPATITIS B VIRUS

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: One Market Plaza, Steuart Street tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,950
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US
; FILING DATE:
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pameleee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14740-2-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 845 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-416-950-11

```

Alignment Scores:

```

Pred. No.: 1,25e-243 Length: 845
Score: 2759.00 Matches: 516
Percent Similarity: 97.23% Conservative: 10
Best Local Similarity: 95.38% Mismatches: 14
Query Match: 46.71% Indels: 2
DB: 1 Gaps: 0

```

US-10-761-006a-1 (1-3215) x US-08-416-950-11 (1-845)

```

QY 1 CTCACAACTTCACCAAGCTCTGTAGATCCAGGGGTGAGGGGCTATATTTTCTGTC 60
Db 306 LeuHisAsnIleProProAsnSerAlaArgSerGlnSerGlyProValPheSerCys 325
QY 61 TGTGGCTCCAGTCCGGACAGTAAACCTGTTCGACTACTGCTCTCCCATATCGTC 120
Db 326 TrpIrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 345
QY 121 AATCTTTCTCGAGACTGGGACCTCCACCGAATCATGGAGAACACAACTATCGATTCCCT 180
Db 346 AsnLeuLeuGluAspIrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 365
QY 181 AGAACCCCTGCTGCTTACAGCGGGGTTTTCTCGTTGACAAAGATCTCTCAATACC 240
Db 366 ArgThrProAlaArgValThrGlyIleValPheLeuValAspLysAsnProHisAsnThr 385
QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGAGACCCACGGTTTC 299
Db 386 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrArgVal-Se 405
QY 300 CTGGCCAAATTCGAGTCCCAACCTCCCAATCACTCACCAACCTCTGTCTCTCCCAATTT 359
Db 405 rTtpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 425

```

```

QY 360 GTCTCGCTATCGCTCGATGTGTCTCGGGCTTTTATCATATTCCTTCTTCATCTCTGCTGC 419
Db 425 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 445
QY 420 TATGCTCATCTCTTCTTGTGTCTCTGACTTACCAAGGTATGTGTCCCGTTGTCTCTC 479
Db 445 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 465
QY 480 TATCTCCAGGAACATCAACCAACGACGCGGGCCATGCAAGAGCTCGACAGCTCTGCTCT 539
Db 465 rAsnSerArgIleAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 485
QY 540 AAGAAACTCTAGCTTTCTCTTGTGTCTGTACAAAACCTTCGGAGCGGAACCTGACATT 599
Db 485 rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 505
QY 600 GTATTCCTCCATCCATCATCTCGGGCTTCGCAAGATTCTATGGGAGTGGGCTCTAGTCC 659
Db 505 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 525
QY 660 GTTTCCTCTGGCTCAGTTTACTAGTCCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCA 719
Db 525 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaAlaPheProHi 545
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTGTATTTGGGGGGGGAAGTCTGTACAACTCT 779
Db 545 sCysLeuAlaAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 565
QY 780 TGAGTCCCTTTTACCTCTTATTACCAATTTCTCTTGTCTTGTGGGTATACATTTAAACCC 839
Db 565 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 585
QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACTTCATGGGATATGTAAATCGAAGTTG 899
Db 585 oAsnLysThrLysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 605
QY 900 GGGTACTTTTACCGCAGGAACATATGTACTTAAACTCAAGCAATGTTTTCGAAAACCTGCC 959
Db 605 pGlyThrIleProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuPr 625
QY 960 TGTAAATAGACCTATTGATTGGAAAGTATGTCAAGAAATGTGGGTCTTTTGGGCTTTGC 1019
Db 625 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 645
QY 1020 TGCCCCCTTTTACAAATGTGGCTATCCTCGCTTGATGCTCTTTATATGCATGTATACAATC 1079
Db 645 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 665
QY 1080 TAAGCAGGCTTTCACCTTTCTCGCAACTTACAAGGCTTTCTGTGTAAACAATATCTGAA 1139
Db 665 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 685
QY 1140 CTTTTACCCCGTTGCGCGCAACGGTCTCTCGCAAGTGTGTTCGTCGACCAACCC 1199
Db 685 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 705
QY 1200 CACTGGATCGGCTTGGCCATAGGCCATCAGGCGCATGGGTGGAACTTTCTGCTCTCTCT 1259
Db 705 oThrGlyTyrGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 725
QY 1260 GCGGATCCATCTAGCGGAACCTCTAGCAGCTTGTGTGTCTCGCAGCGGCTCTGGAGCAA 1319
Db 725 uProfileHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 745
QY 1320 ACTTATCGGAACCGCAACTCTGTCTCTCGAAATATACACCTCTCTTTCATGGCT 1379
Db 745 sleuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 765
QY 1380 GTTAGGGTGTGTGCAACTGGATCTCGCGGGAGCTCTTTGTCTAGTCCCGCTGGC 1439
Db 765 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 785
QY 1440 GCTGAATCCCGCGGACGACCCGCTCTCGGGCGGCTTACCGTCCCTCTTCTTCA 1499

```

Db 785 aLeuAsnProHlaAspProSerArgGlyArgLeuGlyLeuTyArgProLeuLeuAr 805
Qy 1500 TCTCCGCTCCGGCCGACCGGCGCAGCTCTCTTTACGCGGTCTCCCGGTATGTGCC 1559
Db 805 gLeuProPheArgProThrThrGlyA:GThrSerLeuTyAlaValSerProSerValPr 825
Qy 1560 TTCTCANTCTCCGACCGGTGCACTTCGCTTCACTCTCAGCTCCGATGGAGACCA 1619
Db 825 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaIatPrArgProPr 845
Qy 1620 G 1620
Db 845 o 845

RESULT 9

US-08-469-830-11
; Sequence 11, Application US/08469830
; Patent No. 5932224
; GENERAL INFORMATION:
; APPLICANT: CHISARI, Francis V.
; TITLE OF INVENTION: PEPTIDES FOR INDUCING CYTOTOXIC T
; TITLE OF INVENTION: LYMPHOCYTE RESPONSES TO HEPATITIS B VIRUS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,830
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Farmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14740-2-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 845 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-469-830-11

Alignment Scores:

Pred. No.: 1-25e-243 Length: 845
Score: 2759.00 Matches: 516
Percent Similarity: 97.23% Conservative: 10
Best Local Similarity: 95.38% Mismatches: 14
Query Match: 46.71% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1 (1-3215) x US-08-469-830-11 (1-845)

Qy 1 CTCACAAACATTCACCAAGCTCTGCTAGATCCAGGGTGAGGGGCTATATTTTCTCTGC 60
Db LeuHisAsnIleProProAsnSerAlaArgSerGlnSerGlnGlyProValPheSerCys 325
Qy 61 TGTGTGCTCCAGTTCGCGAACAGTAACCTGTTCCGACTACTGCTCTCCCATATGCTC 120
Db TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrcysLeuThrHisIleVal 345
Qy 121 AATCTTCTCGAGACTGGGGACCTCCACGAAATGGAGAACACACATCAGGATTCCT 180
Db AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyHisAsnIleArgIlePro 365
Qy 181 AGGACCTCTCTGCTTACAGCGGGGTTTTCTCTGTTGACAAAGATCTCTCAATPACC 240
Db ArgThrProHlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 385
Qy 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGCACCCAGTGCTC 299
Db ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrArgVal-Se 405
Qy 300 CTGGCCAAATTCGAGTCCCGACCTCCCAATCCTCACTACCAACCTCTGTCTCTCAATTT 359
Db rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 425
Qy 360 GTCCTGGCTATCGCTGGATGTCTCGCGGGTTTTATCATATTCCTTCTCATCTCTGCTGC 419
Db uSerTrpLeuSerLeuAspValSerAlaAlaPheTyHisIleProLeuHisProAlaAl 445
Qy 420 TATGCTCATCTTCTGTTGTTCTCTGACTACCAAGGTATGTTGCCGCTTTGCTCTCTC 479
Db aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyValAlaArgLeuSerSe 465
Qy 480 TACTTCCAGGAACATCAACACGACGCGGCGCATCGACACCTGCACGACTCTCTGCTC 539
Db rAsnSerArgIleAsnTyrcysGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 485
Qy 540 AAGGAAACTCTACGTTTCCCTCTGTTGCTGTACAAAACCTTCGAGCGGAACTGCATTT 599
Db rArgAsnLeuTyValSerLeuLeuLeuLeuTyLysThrPheGlyArgLysLeuHisLe 505
Qy 600 GTATTTCCCATCCCATCATCTCGGCTTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
Db uTyrcysHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 525
Qy 660 GTTTCCTCTGCTCAGTTTACTAGTGCATTTGTTCACTGCTTCGTAGGCTTTCCGCCA 719
Db oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 545
Qy 720 CTGTTTCTGCTTTCAGTTATATGATGTGTTTGGGGCGAAGTCTGTACCAACATCT 779
Db scysLeuAlaPheSerTyrcysMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 565
Qy 780 TGAGTCCCTTTTACCTCTATTACCAATTTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
Db uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 585
Qy 840 TAATAAACCAACGTTGGGCTACTCCCTTAACCTTCATGGGATATGTAATTTGGAAGTTG 899
Db oAsnLysThrLysArgTrpGlyTyrcysLeuAsnPheMetGlyTyrcysValIleGlySerTr 605
Qy 900 GGGTACTTTTACCGCAGGAACATATTGTACTAAAACTCAAGCAATGTTTTTCGAAAACTGCC 959
Db pGlyThrIleProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuPr 625
Qy 960 TGTAATAGACCTATTGATTGGAAAGTATGTCAAAGAAATTTGGGTCTTTTGGGCTTTGCT 1019
Db oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 645
Qy 1020 TGCCCTTTTACAAATGTGGCTATCTCGCTTGTATGCTTTATATGATGATATACAAATC 1079
Db aAlaProPheThrGlnCysGlyTyrcysProAlaLeuMetProLeuTyrcysIleGlnSe 665

QY 1080 TAAGCAGGCTTTTCTGCGCAACTTACAGGGCTTTTCTGTGTAAACAATATCTGAA 1139
Db 665 rLysGlnAlaPheThrPheSerProThrTyrlYsAlaPheLeuCysLysGlnTyrlLeuAs 685
QY 1140 CTTTATCCCGTTCGCGGCAACGGTCCGCTCTCTGCAAGTGTTCGTGACCAACCCC 1199
Db 685 nLeuTyrlProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 705
QY 1200 CACTGGATGGGCTTGCCCATAGCCATCAGCCATGCGTGGAACTTTCTGGCTCTCT 1259
Db 705 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 725
QY 1260 GCCGATCCATCTGCGGAACCTCTAGCAGCTTGTTCGTCGAGCGGTCTGGAGCAAA 1319
Db 725 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 745
QY 1320 ACTTATCGGAACCGACAACTCTGTGTCCTCTCGGAATATACCTCTTTCATGGCT 1379
Db 745 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrlThrSerPheProIrlPle 765
QY 1380 GCTAGGCTGTGCTGCCAACTCGATCTCGCGGGACGTCCTTTGTCTACGTCGCCGCGC 1439
Db 765 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrlValProSerAl 785
QY 1440 GCTGAATCCCGCGGACGACCCGCTCTCGGGCCGTTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db 785 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrlArgProLeuLeuAr 805
QY 1500 TCTGCGGTTCCGCGCACCAACGCGGCGACCTCTCTTTAGCGGCTCTCCCGPATGCGC 1559
Db 805 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrlAlaValSerProSerValPr 825
QY 1560 TTCTCATCTGCCGACCGTGTGCACTTCGCTTCACCTCTGCACTGCGCATGGAGACCAACC 1619
Db 825 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 845
QY 1620 G 1620
Db 845 o 845
RESULT 10
US-08-591-502B-11
; Sequence 11, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chiesari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685

; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 845 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..845
; OTHER INFORMATION: /product= "OTHER"
; /note= "Xaa = any amino acid
; (<50% consensus)"
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-08-591-502B-11
Alignment Scores:
Pred. No.: 5.56e-242 Length: 845
Score: 2741.00 Matches: 514
Percent Similarity: 96.30% Conservative: 7
Best Local Similarity: 95.01% Mismatches: 19
Query Match: 46.40% Indels: 2
DB: Gaps: 0
US-10-761-006A-1 (1-3215) x US-08-591-502B-11 (1-845)
QY 1 CTCACAACATTTCCACCAAGCTCTGCTAGATCCCGGGTGAGGGGCTATATTTTCTCTGC 60
Db 306 LeuHisAsn***ProProAsnSerAlaArgSerGlnSerGlnGlyProValPheSerCys 325
QY 61 TGGTGGCTCCAGTTCCTCGGAACAGTAAACCTCTGTCGACTACTGCTCTCTCCCATATCGTC 120
Db 326 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyCysLeu**HisIleVal 345
QY 121 ATCTTCTCAGGACTGGGGACCCCTGCACCGAACATGGAGAACACAAATCAGGATTCCT 180
Db 346 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHis***IleArgIlePro 365
QY 181 AGACCCCTGCTCGTGTACAGCGGGGTTTTCTCGTTGACAAGAAATCCTCAATAACC 240
Db 366 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 385
QY 241 GCAGAGTCTAGACTCTG-CTGGACTTCTCTCAATTTTCTAGGGGGAGCACCCACGTGTTTC 299
Db 386 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGly***ThrArgVal-Se 405
QY 300 CTGGCCAAAATTCGAGTCCCAACCTCCAATCACTCACCACCTCTTCTCTCCCAATTT 359
Db 405 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 425
QY 360 GTCTCGGCTATCGTGGATGTCTGCGCGGTTTTATCATATTTCTCTTTCATCTCTCTGC 419
Db 425 userTrpLeuSerLeuAspValSerAlaAlaPheTyrlHisIleProLeuHisProAlaAl 445
QY 420 TATGCTCATCTTCTTGTGTTCTTCTGACTTACCAAGGTATGTTGCCGTTTGTCTCTC 479
Db 445 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrlValAlaArgLeuSerSe 465
QY 480 TACTTCCAGGAACATCAACACGACGCGGGCCATGCAAGACCTGCACGACTCTCTGCTC 539
Db 465 rAsnSerArgIleIleAsn***GlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 485
QY 540 AAGGAAACTCTACGTTTCCCTCTTGTGTGTGTACAAAACCTTCGGACGGAACCTGACATT 599
Db 485 rArgAsnLeuTyrlValSerLeuLeuLeuLeuTyrlsThrPheGlyArgLysLeuHisLe 505

```

600 GTATTCCCATCCATCTCTGGGCTTTCCAGATTCTCTATGGAGTGGCCTCAGTCC 659
605 uTySerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 525
660 GTTTCCTCTGGCTCAGTTTACTAGTGCATTTGTTCAGTGTCTCGTAGGCTTTCCCCCA 719
665 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 545
720 CTGTTTGGCTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 779
745 sCysLeuAlaPheSerTyrrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 565
780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGTATACATTTAAACCC 839
785 uGluSerLeuPheThrAlaValThrAsnPheLeuSerLeuGlyIleHisLeuAsnPr 585
840 TAAATAAACCAAGCTTGGGCTACTCCCTTAACTTTCATGGATATGTAATGGAAGTTG 899
845 oAsnLysThrLysArgTrpGlyTyrrSerLeuAsnPheMetGlyTyrrValIleGlySerTr 605
900 GGGTACTTTACCGAGGAACATATGTTACTTAACTAACTACAGCAATGTTTTCGAAAACGTC 959
905 pGlyThrLeuProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuPr 625
960 TGTAAATAGACTATTGATTTGGAAGATATGTCGAAGATTTGGGCTTTTGGGCTTTGC 1019
965 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 645
1020 TGCCCTTTTACAAATGTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079
1045 aAlaProPheThrGlnCysGlyTyrrProAlaLeuMetProLeuTyrrAlaCysIleGlnSe 665
1080 TAAGCAGGCTTTTCACTTTCTCGCAACTTACAAGGCTTTCTGTGTAAACAATATCTGAA 1139
1085 rLysGlnAlaPheThrPheSerProThrTyrrLysAlaPheLeuCysLysGlnTyrrLeuAs 685
1140 CCTTTACCGGCTGCGGCAACGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCT 1199
685 nLeuTyrrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 705
1200 CACTGATGGGCTTGGGCATAGGCATCAGCGATGCTCGAACCTTTCTGGCTCTCTCT 1259
705 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 725
1260 GCGATCCATATCGCGGAACCTCTAGCAGCTTGTGCTCGCAGCGGCTCGAGCAAA 1319
725 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLy 745
1320 ACTTATCGGAACCGCAACTCTGTGCTCTCTCGGAATACACCTCTTCCATGGCT 1379
745 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrrThrSerPheProTrpLe 765
1380 GCTAGGGTGTGTCGCAACTGATCTCGCGGAGCTCTTTGTCTACGTCCTCGCTCGGC 1439
765 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrrValProSerAl 785
1440 GCTGAATCCCGGAGCAGCCGCTCTCGGGGCGGTTTGGGCTCTACCGTCCCTTCTTCA 1499
785 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrrArgProLeuLeuAr 805
1500 TCTGCGGCTCGCGGACGACGCGGCGGACCTCTCTTTACCGGCTCTCCGCTATGTC 1559
805 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrrAlaValSerProSerValPr 825
1560 TTCTCATCTGCGGACCGTGTGCACTCTGCTTCACTCTGCACTGCGATGAGACACC 1619
825 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 845
1620 G 1620
845 O 845

```

```

RESULT 11
US-08-591-502B-51
; Sequence 51, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 842 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-08-591-502B-51
Alignment Scores:
Pred. No.: 2,19e-241 Length: 842
Score: 2734.50 Matches: 517
Percent Similarity: 96.86% Conservative: 7
Best Local Similarity: 95.56% Mismatches: 15
Query Match: 46.29% Indels: 3
DB: 2 Gaps: 1
US-10-761-006A-1 (1-3215) x US-08-591-502B-51 (1-842)
Qy 1 CTCACCAACATTCACCAAGCTCTGTAGATCCACGGGTGAGGGCCCTATATTTCTCTGC 60
Db 304 LeuHisIleSerProSerProAlaArgSerGlnSerGluGlyProllePheSerSer 323
Qy 61 TGTGTGGCTCCAGTTTCCGAAACAGTAAACCCCTGTTCCGACTACTCCCTCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrrCysLeuThrHisIleVal 343
Qy 121 AATCTTCTCGAGGACTGGGACCTCGACCGAACATCGAGAACACACATCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363

```

181 AGGACCCCTCGTGTACAGCGGGTTTTCTCGTTGACAGAACTCCTCAATACC 240
 364 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 383
 241 GCAGAGTCTAGACTCG-GTGGACTCTCTCAATTTTCTAGGGGGAGCACCCACGTGTTTC 239
 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
 300 CTGGCCAAATTCGAGTCCCAACCTCCAACTCACTCACCAACTCTGTGTCCTCAATTT 359
 403 rTTPProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 423
 360 GTCCCTGCATCGCTGATGCTGCGGGGTTTTATCATATTCCTCTTCATCTCGTGC 419
 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAla 443
 420 TATGCTCTCATCTCTTCTGTTGTTCTCTGACTACCAAGGTATGTTGCCCGTTGCTCT 479
 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
 480 TACTTCAGGAACATCAACCAACGACGCGGCATGCAAGACTGCAAGACTGCTGCTC 539
 463 rThrSerArgAsnIleAsnHisGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
 540 AAGGAACTCTACGTTTCCCTCTGTTGCTGTACAAACCTTCGGACGGAACTGCACATT 599
 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
 600 GTATTCCTCATCCATCATCTCGGCTTTCGCAAGATTCTTATGGAGTGGGCTCAGTCC 659
 503 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyGlyLeuSerPr 523
 660 GTTTCCTCGCTCAGTTTACTAGTGCATTTGTTAGTGGTTCGATGGCTTTCCTCCCA 719
 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
 720 CTGTTTGGCTTTCAGTTATATGATGATGATGTTGTTGGGCGAAGTCTGTACAACTCT 779
 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGTTGTTGGGTATACATTTAAACCC 839
 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
 840 TAATAAACCAACGTTGGGCTACTCCCTTAATCTCATGGATATGTAATTTGGAAGTTG 899
 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
 900 GGGTACTTTACCGCAGGAATATTGTACTAAACTCAAGCAATGTTTTCGAAACTGCC 959
 603 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 623
 960 TGTAATATAGACCTATTGATTGGAAGTATGTCAAGAAATTTGGGTCTTTGGGCTTTC 1019
 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
 1020 TGCCCCCTTTTACAAATGTGGCTATCTCGCTGATGCTGATGCTTTATGATCATGTATACAATC 1079
 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
 1080 TAAGCAGCTTTCATCTTCTCGCAACTTACAGGCCCTTTCTGTGTAAACAATCTGAA 1139
 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
 1140 CTTTACCCCGTTCCCGGCAACGGTCTGCTGCAAGTGTGTTCTGACGCAACCCCC 1199
 683 sLeuTyrProValAlaArg--ArgThrAlaLeuCysGlnValPheAlaAspAlaThrPr 702
 1200 CACTGGATGGGCTGTGGCCATAGCCATAGCCGATGCGTGGAACTTTCTGGCTCCTCT 1259
 702 oThrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 722
 1260 GCCGATCCATACGCGGAACCTCTAGCAGCTTGTGTTTGTCTCGCAGCGGCTCGAGCAA 1319

722 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 742
 1320 ACTTATCGGAACGCAACTCTGTTGCTCTCTCTCGGAATAACACTCTCTTCCATGGCT 1379
 742 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 762
 1380 GGTAGGTGTGTGCTCCAACTGATCTCGCGGGAGCTCTTTGTCTAGTCCCGTCCGCG 1439
 762 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrTyrPheValTyrValProSerAl 782
 1440 GGTGAATCCCGCGGACGACCCGCTCTCGGGCGGCTTTGGGGCTCTACCGTCCCCCTTTCA 1499
 782 aLeuAsnProAlaAspAspProSerArgLysArgLeuGlyLeuIleArgProLeuLeuHi 802
 1500 TCTGCCGTTTCCGGCGACCAACGGGGCGACCTCTCTTTTACGGCGTCTCCCGTATGTGCC 1559
 802 sLeuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 822
 1560 TTCTCATCTCGCGGACCGTGTGCATCTGCTTTCACCTCTGACGTCGATGGAGACCACC 1619
 822 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 842
 1620 G 1620
 842 o 842
 RESULT 12
 US-08-591-502B-50
 ; Sequence 50, Application US/08591502B
 ; Patent No. 6607727
 ; GENERAL INFORMATION:
 ; APPLICANT: Chisari, Francis V.
 ; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
 ; Lymphocyte Responses to Hepatitis B Virus
 ; NUMBER OF SEQUENCES: 99
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/591,502B
 ; FILING DATE: 20-May-1996
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/749,540
 ; FILING DATE: 26-AUG-1991
 ; APPLICATION NUMBER: US 07/935,898
 ; FILING DATE: 26-AUG-1992
 ; APPLICATION NUMBER: US 08/100,870
 ; FILING DATE: 02-AUG-1993
 ; APPLICATION NUMBER: WO PCT/US94/08685
 ; FILING DATE: 01-AUG-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Weber, Ellen Lauver
 ; REGISTRATION NUMBER: 32,762
 ; REFERENCE/DOCKET NUMBER: 014740-000230US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 576-0200
 ; TELEFAX: (415) 576-0300
 ; INFORMATION FOR SEQ ID NO: 50:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 842 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: <Unknown>

;
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 50:
 ; US-08-591-502B-50
 Alignment Scores:
 Pred. No.: 6, 28e-241 Length: 842
 Score: 2739.50 Matches: 516
 Percent Similarity: 96.67% Conservative: 7
 Best Local Similarity: 95.38% Mismatches: 16
 Query Match: 46.21% Indels: 3
 DB: 2 Gaps: 1
 US-10-761-006a-1 (1-3215) x US-08-591-502B-50 (1-842)
 QY 1 CTCACAACTTCACCAAGCTCTGCTAGATCCAGGGTGGGGCCCTATATTTCTCTGC 60
 Db 1
 QY 304 LeuHisHisSerProAlaArgSerGlnSerGluGlyProIlePheSerSer 323
 Db 304
 QY 61 TGTGGCTCCAGTTCGGGAACAGTAACCCCTGTTCCGACTACTCCCTCCCATATCGTC 120
 Db 61
 QY 324 TrpTrpLeuGlnPheArgSerLysProCysCysAspTyrCysLeuThrHisIleVal 343
 Db 324
 QY 121 AATCTTCTCAGGACTGGGACCTCGCACCGAATCGAGAACACACATCAGGATTCCT 180
 Db 121
 QY 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
 Db 344
 QY 181 AGGACCTCTCTGCTGTATACAGGGGGGTTTTCTCTGTGACAAAGATCTCTCAATACC 240
 Db 181
 QY 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
 Db 364
 QY 241 GCAGAGTCTAGACTCTGCTGCTCAATTTCTAGGGGAGGACCCAGCTGTC 299
 Db 241
 QY 384 ThrGluSerArgLeuValAlaAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
 Db 384
 QY 300 CTGGCCAAATTCGAGTCCCACTCAATCACTACCAACTCTTGTCTCTCAATTT 359
 Db 300
 QY 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
 Db 403
 QY 360 GTCCTGGCTATCGTGGATGTCTGCGCGGTTTTTATCATATTCCTCTCTCAATCTGCTGC 419
 Db 360
 QY 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
 Db 423
 QY 420 TATGCTCATCTCTGTTGTTCTCTGACTACCAAGGTATGTTGCCCGTTGTCCTC 479
 Db 420
 QY 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
 Db 443
 QY 480 TACTTCAGGACATCAACCCAGCAGCGGCCATGCAAGACTGACGACTCTCTGCTC 539
 Db 480
 QY 463 rThrSerArgAsnIleAsnHisGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
 Db 463
 QY 540 AAGGAACTCTACGTTTCCCTCTCTGCTGTACAAAACCTTCGGACGGAACCTGCACCT 599
 Db 540
 QY 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
 Db 483
 QY 600 GTATTCATCCATCATCTGGGCTTTCGCAAGATTCTCTATGGAGTGGGCTCAGTCC 659
 Db 600
 QY 503 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyGlyGlyLeuSerPr 523
 Db 503
 QY 660 GTTCTCTGGCTCAGTTTACTAGTGCATTTGTTTCAGTGGTTCTGAGGGCTTTCGCCCA 719
 Db 660
 QY 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHI 543
 Db 523
 QY 720 CTGTTGGCTTTCAGTTATATGATGATGTGTATTTGGGGGCGAAGTCTGTACACATCT 779
 Db 720
 QY 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
 Db 543
 QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACC 839
 Db 780
 QY 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
 Db 563
 QY 840 TAAATAAACCAACAGTTGGGGCTACTCCCTTAATCTCATGGATATGTAATTTGGAAGTTG 899
 Db 840

Db 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
 QY 900 GGTACTTTTACCGCAGCAATATTTACTATAAATCTCAAGCAATGTTTTCGAAAACCTGCC 959
 Db 603 pGlyThrLeuProGlnGluHisIleValLeuLysIleValGlnCysPheArgLysLeuPr 623
 QY 960 TGTAAATAGACCTATTGATTGAAAGTATGTCAAAGATTGGGCTCTTTGGGCTTTGC 1019
 Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
 QY 1020 TGGCCCTTTTACCAATGTGGCTATCTCTGCTGTATGATGCTTTATATGATGATATCAATC 1079
 Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
 QY 1080 TAAAGAGGCTTTTCACTTTCTCGCAACTTACAAAGGCTTTTCTGTAAACAAATATCTGAA 1139
 Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
 QY 1140 CTTTATCCCGGTTTGGCCGCAACGGTCTCTGCTGCAAGTGTCTGCTGACGCAACCCC 1199
 Db 683 sLeuTyrProValAlaArg---ArgThrAlaLeuCysGlnValPheAlaAspAlaThrPr 702
 QY 1200 CACTGGATGGGCTTGGCCATAGCCATAGCCATAGCCATAGCCATAGCCATAGCCATAGCC 1259
 Db 702 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 722
 QY 1260 GCCGATCCATCTACGCAACTCTCTGCTGCTCTCTCGGAAATACACCTCTTCCATGGCT 1319
 Db 722 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 742
 QY 1320 ACTTATCGGAACCGCAACTCTGTTGCTCTCTCGGAAATACACCTCTTCCATGGCT 1379
 Db 742 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 762
 QY 1380 GTTAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1439
 Db 762 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrTyrPheValTyrValProSerAl 782
 QY 1440 GCTGAATCCCGGACGACCGCTCTCGGGCGCGTTCGGGGCTCTACCGTCCCTCTTCTCA 1499
 Db 782 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuIleArgProLeuLeuHi 802
 QY 1500 TCTGCTGCTTCCGGCGCACCGCTCTCTTTAGCGGTCTCTCCCGTATGTGCC 1559
 Db 802 sLeuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 822
 QY 1560 TTTCTATCTGCGGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1619
 Db 822 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaIleTrpArgProPr 842
 QY 1620 G 1620
 Db 842 o 842
 RESULT 13
 ; US-08-591-502B-61
 ; Sequence 61, Application US/08591502B
 ; Patent No. 660727
 ; GENERAL INFORMATION:
 ; APPLICANT: Chisari, Francis V.
 ; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
 ; Lymphocyte Responses to Hepatitis B Virus
 ; NUMBER OF SEQUENCES: 99
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591.502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994

ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 61:
SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-08-591-502B-61
SEQUENCE CHARACTERISTICS:
LENGTH: 832 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-08-591-502B-61

Alignment Scores:
Pred. No.: 7,84e-234 Length: 832
Score: 2652.00 Matches: 494
Percent Similarity: 95.38% Conservative: 22
Best Local Similarity: 91.31% Mismatches: 24
Query Match: 44.90% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x US-08-591-502B-61 (1-832)

QY	1	CTCCAGCAATTCACCAAGCTCTGTAGTACCCAGGCTGAGGGCTATATTTCTCTGC	60
Db	293	LeuHisAsnLeuProAsnSerAlaArgSerGlnSerGluArgProValPheProCys	312
QY	61	TGTTGGCTCCAGTTCGGAAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTC	120
Db	313	TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal	332
QY	121	AATCTTCTCGAGACTGGGACCTCGACCGAATCATGGAGAACACATCATGATTCCT	180
Db	333	AsnLeuLeuGluAspTrpGlyProCysAlaGluHisGlyGluHisIleArgIlePro	352
QY	181	AGGACCTCTCGTGTACAGCGGGGTTTCTGTTGACAGAGATCCTCATATACC	240
Db	353	ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr	372
QY	241	GCAGAGTCTAGACTCTG-GTGGACTTCTCAATTTTCTAGGGGAGACCCACGCTGTTCC	299
Db	373	AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnTyrArgVal-Se	392
QY	300	CTGGCCAAATTCGAGTCCCAACTCCAACTCACTCAACCAACTCTGTGCTCCCAATTT	359
Db	392	rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe	412
QY	360	GTCTCGCTATCGGTGTCGCGGGTGTATATTCATTTCTCTTCATCTCTCTGC	419
Db	412	uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl	432
QY	420	TATGCTCATCTCTTGTGTTGTTCTCTGAGTACCAAGGTATGTTGCGCGTGTGCTCTC	479
Db	432	aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe	452

QY	480	TACTTCCAGGAACATCAACACGAGCAGCGGGCCATCGAAGACCTGCACGACTCTCTGCTC	539
Db	452	rAsnSerArgIlePheAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe	472
QY	540	AAGGAACTCTACGTTTCCCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	599
Db	472	rArgAsnLeuTyrValSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	492
QY	600	GTATTCCCATCCCATCATCTCTGGGCTTTCGCAAGATTCTATGGGAGTGGGCTCAGTCC	659
Db	492	uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr	512
QY	660	GTTTCTCTCGGCTCAGTTTACTAGTGCATTTTGTTCAGTGGTTCGTAGGGCTTCCCTCC	719
Db	512	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi	532
QY	720	CTGTTTGGCTTTTACGTTATATGATGATGATGATGATGATGATGATGATGATGATGATG	779
Db	532	sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe	552
QY	780	TGAGTCCCTTTTACCTCTATTACCAATTTTCTTTTGTCTTTTGGGTATACATTTAAACCC	839
Db	552	uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr	572
QY	840	TAATAAAACCAACGTTGGGCTACTCCCTTAACCTTACCTGATGATGATGATGATGATG	899
Db	572	oAsnLysThrLysArgTrpGlyTyrSerLeuHisPheMetGlyTyrValIleGlyCysTy	592
QY	900	GGGTACTTTTACCGCAGGAACATATTGTACTAAACCTCAAGCAATGTTTTCGAAACCTGC	959
Db	592	rGlySerLeuProGlnAspHisIleIleGlnLysIleLysGluCysPheArgLysLeuPr	612
QY	960	TGTAATAGACCTATTGATTGGAAGATGATGTCAAGAGATTGTGGGTCTTTTGGCTTTGC	1019
Db	612	oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl	632
QY	1020	TGCCCCCTTTTACCAATGTCCTTCTCCCTTGTATGTCCTTGTATGATGATGATGATG	1079
Db	632	aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe	652
QY	1080	TAAGCAGGCTTTTCACTTTCTCCCAACTTACAAAGGCTTTCTGTGTAAACAATATCTGAA	1139
Db	652	rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs	672
QY	1140	CCTTTACCCCGTTGCGGCAACGGTCCGCTCTCTGCCAAGTCTTGTGTCGACCAACCCC	1199
Db	672	nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr	692
QY	1200	CACTGATGGGCTTGGCCATAGGCCATCAGGCATGGCTGGAACTTTCTGGCTCTCTCT	1259
Db	692	oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaArgLe	712
QY	1260	GCGATCCATATCGGGAACCTCTAGCAGCTGTTGTTGTCGAGCGGCTCTGGAGCAAA	1319
Db	712	uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs	732
QY	1320	ACTTATCGGAACCGCAACTCTGTTGTCTCTCTCGAAATACACCTCTTTTCCATGGCT	1379
Db	732	nIleLeuGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerTyrProTrpLe	752
QY	1380	GCTAGGCTGTCTGCCAATGATCTCTGCGCGGAGCTCTTTGTCTAGTCCCGTCCGCG	1439
Db	752	uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl	772
QY	1440	GCTGAATCCCGGAGCAGCCCTCTCGGGCGGCTTTGGGGCTCTACCGTCCCTTCTTCA	1499
Db	772	aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuSerArgProLeuLeuAr	792
QY	1500	TCTGCGGTTCCGCGGAGCAGCCCGGCGCACCTCTCTTTTACGGGCTTCCCGTATGCTC	1559
Db	792	gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr	812

1560 TTCTCTCTCCGACGCTGTCACCTTCCTCTCACCTCTCGACGTGCGATGGAGACACC 1619
812 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832

1620 G 1620

832 o 832

RESULT 14
US-08-591-502B-62
Sequence 62, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus

NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA

ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Leuwer
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-00023005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 832 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 62:
US-08-591-502B-62

Alignment Scores:
Pred. No.: 2,78e-233 Length: 832
Score: 2646.00 Matches: 495
Percent Similarity: 95.01% Conservative: 19
Best Local Similarity: 91.50% Mismatches: 26
Query Match: 44.79% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x US-08-591-502B-62 (1-832)

1 CTCCACAACTTCCACCAAGCTCTGCTAGATCCCAAGGTGAGGGCGCTATATTTCTCTGC 60
233 LeuHisAsnLeuProProAsnSerAlaArgSerGlnSerGluArgProValPheProCys 312

61 TGGTGGCTCCAGTTCGGGAACAGTAAACCCCTGTTCGACTACTGCTCTCTCCCATATCGTC 120
313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerLeuLeuVal 332
121 AATCTTCTCAGGACTGGGGACCCCTGCACCGAATGGAGACACACAACTCCTCAATATCCT 180
333 AsnLeuLeuGluAspTrpGlyProCysAlaGluHisGlyGluHisIleArgIlePro 352
181 AGGACCCCTGCTGCTGTATACAGCGGGGTTTTCTCGTTGACAGAATCTCTCAATATCCT 240
353 ArgThrProSerArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 372
241 GCAGACTCTAGACTCTG-GTGGAATCTCTCAATTTCTAGGGGGAGACCCACAGTGTTC 299
373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnTyrArgVal-Se 392
300 CTGGCCAAATTCGCAGTCCCAACCTCCCAATCACCACCAACCTCTGTCTCTCCCAATTT 359
392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 412
360 GTCCTGCTATCGTGTGTCTGCGGGTTTTATCATATTCCTCTCTTCTCATCTCTGCTGC 419
412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 432
420 TATGCTCATCTCTTGTGTTGCTTCTTCTGACTTACCAAGGTATGTGCTGCTGCTCTCCT 479
432 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 452
480 TACTTCCAGAAACATCAACACGAGCGGGCCATGCAAGACCTGCACACCTCTGCTCTCCTC 539
452 rAsnSerArgIleLeuAsnHisGlnHisGlyThrMetProAsnLeuHisAspSerCysSe 472
540 AAGGAACTCTAGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 599
472 rArgAsnLeuTyrValSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLe 492
600 GTATTCCTCCATCCATCATCTCTGCTTCTCGCAAGATTCCTATGGAGTGGGCTCTCAGTTC 659
492 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512
660 GTTCTCTCTGCTCAGTTACTAGTGCATTTGTTTCTAGTGGTTCGTAGGCTTTCCCTCCA 719
512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 532
720 CTGTTTGGCTTTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGATG 779
532 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 552
780 TGAGTCCCTTTTACCTCTATTACCAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 839
552 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 572
840 TAATAAAACCAACGTTGGGCTACTCCCTTAACCTTACCTTATGATGATATGATGATGATG 899
572 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlyCysTy 592
900 GGGTACTTTTACCGCAGGAACATATGTTACTTAAACTCAAGCAATGTTTTCGAAACTGCC 959
592 rGlySerLeuProGlnGluHisIleIleGlnLysIleLeuGlyCysPheArgLysLeuPr 612
960 TGTAAATAGACCTATTGATTGGAAGATGATGTCAAGAAATTTGGGTCTTTTGGGCTTTGCG 1019
612 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
1020 TGCCCCCTTTTACAAATGTGGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079
632 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 652
1080 TAAGCAGCTTTTCACTTTCTCGCCCAACTTACAAGSCCTTTCTGTGTAAACAATATCTCAA 1139
652 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 672
1140 CTTTATACCCCTGTCGCCGCAACGGTCTCGGTCTCTGCAAGTGTGTTGTGACGCAACCCC 1199

```
Db 672 nLeuTyProValAlaGlyGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
QY 1200 CACTGATGGGCTTGCCATAGCCATCAGCCATGCTGGAACTTTCTGGCTCTCT 1259
Db 692 oThrGlyTrpGlyLeuAlaMetGlyHisGlnArgMetArgGlyThrPheSerAlaProLe 712
QY 1260 GCCATCCATCTACTGCGGAACTCCTAGCAGCTTGTGCTCGCAGCGGCTCGAGCAAA 1319
Db 712 uProIleHisThAlaGlnLeuLeuAlaLaCysPheAlaArgSerArgSerGlyAlaAs 732
QY 1320 ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGGAATACACCTCTCTTCATGGCT 1379
Db 732 nIleIleGlyThrAspAsnSerValValLeuSerArgLySerThrSerPheProTrie 752
QY 1380 GCTAGGCTGTGTCGCAACTGTGATCTCGCGGAGCTCTTTGTCTACGTCGCCGCGG 1439
Db 752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAl 772
QY 1440 GCTGAATCCCGCGGAGACCGCTCTCGGGCCGTTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db 772 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuSerArgProLeuLeuAr 792
QY 1500 TCTGCGCTTCCGCGCCAGCACCGGGCGCACTCTCTTTAGCGGTCTCCCGGTATGTGC 1559
Db 792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyAlaAspSerProSerValPr 812
QY 1560 TTCTATCTGCGGACGGTGTCACCTTCGCTTCCCTCTGACGTGCGATGGAGACCACC 1619
Db 812 oSerHisLeuProAspLeuValHisPheAlaSerProLeuHisValAlaAlaTrpArgPro 832
QY 1620 G 1620
Db 832 o 832

RESULT 15
US-08-591-502B-56
; Sequence 56, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
```

```
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 56:
US-08-591-502B-56

Alignment Scores: 3.46e-233 Length: 843
Pred. No.: 2645.00 Matches: 493
Score: 95.19% Conservative: 22
Best Local Similarity: 91.13% Mismatches: 25
Query Match: 44.78% Indels: 2
DB: Gaps: 0

US-10-761-006A-1 (1-3215) x US-08-591-502B-56 (1-843)

QY 1 CTCCACAACATTCCACCAAGCTCTGTAGATCCAGGCTGAGGGGCTATATTTTCTCTGC 60
Db 304 LeuHisHisValProProAsnSerSerArgSerGlnSerGlnGlySerValLeuSerCys 323
QY 61 TGGTGGCTCCAGTTCGGGAACAGTAACCCCTGTTCGAGCTACTGCTCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerIysProCysSerGluHisCysLeuPheHisIleVal 343
QY 121 AATCTTCTCAGGACTGGGGACCTGTCACCGAACAATCGGAGAACACAAATCAGGATTCCT 180
Db 344 AsnLeuIleGluAspTrpGlyProCysAlaGluHisGlyGluHisArgIleArgThrPro 363
QY 181 AGGACCCCTGCTCGTGTACAGCGGGGTTTTCTCGTTGACAAAGAACTCCTCACAATACC 240
Db 364 ArgThrProAlaArgValThrGlyValPheLeuValAspIysAsnProHisAsnThr 383
QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGACCCACGCTGTTTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnThrArgVal-Se 403
QY 300 CTGGGCAAAATTCGAGTCCCAACCTCCAATCAGTCAACCAACCTCTCTCTCCCAATTT 359
Db 403 rTrpProIysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAspLe 423
QY 360 GTCTTGGCTATCGCTGGATGTCTGCGCGGTTTTATCATATTTCTCTTTCATCTCTCTGC 419
Db 423 userTrpLeuSerLeuAspValSerAlaAlaPheTyHisLeuProLeuHisProAlaAl 443
QY 420 TATGCTCATCTTCTTGTGTTGTTCTGTGAGTACCAAGTATGTTGCCGTTGTCCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyTrValAlaAlaArgLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACCACGACGCGGGCCATGCAAGACCTGCACGACTCTCTGCTC 539
Db 463 rAsnSerArgIleIleAsnHisGlnHisArgThrMetGlnAsnLeuHisAspSerCysSe 483
QY 540 AAGGAAACTCTACGTTTCCCTCTGTGTGTGTACAAAACCTTCGGACGCGAAATCGCACIT 599
Db 483 rArgAsnLeuTyTrValSerLeuLeuLeuTyTrIysThrTyGlyArgIleuHisIle 503
QY 600 GTATTCCCATCCCATCATCTCTGGGCTTTTCGCAAGATTCCTATGGAGTGGGGCTCAGTCC 659
Db 503 uTySerHisProIleIleLeuGlyPheArgIlysIleProMetGlyValGlyLeuSerPr 523
QY 660 GTTCTCTCGGCTCAGTTTACTAGTGCCATTTGTTGAGTGGTTCGTAGGGCTTTCCCCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
QY 720 CTGTTTGGCTTTTCAGTTATATGATGTGTGTTATTTGGGGGCGAAGTCTGTGTACAACTCT 779
```

Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGTCCTTTGGGTATACATTTAAACCC 839
Db 563 uGluSerLeuTyrAlaAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAACCACAAAGCTTGGGCTACTCCCTTAACCTTCATGGGATATGTAATTTGGAAGTTG 899
Db 583 oGlnLysThrLysArgTrpGlyTyrSerLeuAsnPhemMetGlyTyrValIleGlySerTr 603
QY 900 GGGTACTTTTACCGCAGGAACATATTGTACTAAAACTCAAGCAATGTTTTCGAAAACTGCC 959
Db 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACCTATTGATTGGAAGATATGTCAAGAATTGTGGGCTCTTTTGGGCTTTC 1019
Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCTTTTACACAATGGGCTATCCTCGCTTCATGCTTTATATGTCATGTATACAAATC 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAl 663
QY 1080 TAAGCAGGCTTTTCACTTTCTCGCAACTTACAAGGCTTTCTGTGTAAACAATATCTGAA 1139
Db 663 aLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuThrLysGlnTyrLeuAs 683
QY 1140 CCTTTACCCGTTGCCGGCAACGGTCGGTCTCTGCAAGTGTTCGTCGACGCAACCCC 1199
Db 683 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 703
QY 1200 CACTGGATGGGCTTGGCCATAGCCATCAGCGCATCGCTGGAACCTTTCTGGCTCTCT 1259
Db 703 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValSerProLe 723
QY 1260 GCCGATCCATACCTCGGAACCTCTAGCAGCTTGTGTTTCTCGCAGCCGGTCTGGAGCAAA 1319
Db 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 743
QY 1320 ACTTATCGGAACCGCAAACTCTGTGTCCTCTCTCGGAAATACACCTCTTCCATGGCT 1379
Db 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
QY 1380 GCTAGGTGTCTGCCAATCGATCCTCGCGGACGCTCTTGTCTACGTCCGTCGGC 1439
Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
QY 1440 GCTGAATCCCGGACGACCGCTCTCGGGGCGGTTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuAr 803
QY 1500 TCTGCCGTTCCGGCGCACACGGGCGCACCTCTCTTTACGGGCTCTCCCGTATGTGCC 1559
Db 803 gLeuProTyrArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr 823
QY 1560 TTCTCATCTGCCGACCGCTGTGCTGCTTCCCTCTGACGCTCGCATGGAGACACC 1619
Db 823 oSerArgLeuProAspArgValHisPheAlaSerProLeuHisValAlaIatPargProPr 843
QY 1620 G 1620
Db 843 O 843

